



Universiteit
Leiden
The Netherlands

TGF- β signaling dynamics in epithelial-mesenchymal plasticity of cancer cells

Fan, C.

Citation

Fan, C. (2024, June 26). *TGF- β signaling dynamics in epithelial-mesenchymal plasticity of cancer cells*. Retrieved from <https://hdl.handle.net/1887/3765351>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3765351>

Note: To cite this publication please use the final published version (if applicable).

TGF- β signaling dynamics in epithelial-mesenchymal plasticity of cancer cells

Chuannan Fan

ISBN: 978-94-6496-100-3

Copyright © Chuannan Fan, Leiden, the Netherlands, 2024. All right reserved. No part of this thesis may be reproduced, stored, translated, or transmitted in any form or by any means now or hereafter, electronic or mechanical without prior written permission of the copyright owner.

Cover: concept and design by Chuannan Fan. Modified in situ hybridization picture of A549 cells showing *LITATSI* expression and localization.

Layout: Qian Wang

Printing: GildePrint

TGF- β signaling dynamics in epithelial-mesenchymal plasticity of cancer cells

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op woensdag 26 juni 2024

klokke 11:15 uur

door

Chuannan Fan

geboren te Gaomi, Shandong, China
in 1992

Promotor: Prof. Dr. P. ten Dijke

Co-promotor: Dr. M.A.F.V. Gonçalves

Leden promotiecommissie:

Prof. Dr. C.H. Heldin (Uppsala University)

Prof. Dr. M. Landström (Umeå University)

Prof. Dr. M.J.T.H. Goumans (LUMC)

Prof. Dr. E.H.J. Danen (LACDR, Leiden University)

Dr. H. Mei (LUMC)

Prof. Dr. R. van Doorn (LUMC)

The research presented in this thesis was performed at the Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands. This research was supported by OncoCode Institute base funds, the Cancer Genomics Centre in the Netherlands (CGC.NL), the ZonMW grant (09120012010061) and the China Scholarship Council.

To My Dear Family and Friends

谨以此书献给亲爱的家人及朋友

Contents

Chapter 1	9
General introduction	
Chapter 2	37
LncRNA <i>LITATS1</i> suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation	
Chapter 3	81
The lncRNA <i>LETS1</i> promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism	
Chapter 4	109
OVOL1 inhibits breast cancer cell invasion by enhancing the degradation of TGF- β type I receptor	
Appendix	149
English Summary	150
Nederlandse Samenvatting	152
List of Publications	154
Curriculum Vitae	155
Acknowledgements	156



Chapter 1

General Introduction

1) TGF- β in Cancer Progression: From Tumor Suppressor to Tumor Promoter

Encyclopedia of Cancer, 3rd edition, 2018, Volume 3

2) Long non-coding RNAs in TGF- β signaling and EMT

3) Scope of this thesis



1) TGF- β in Cancer Progression: From Tumor Suppressor to Tumor Promoter

Chuannan Fan[#], Jing Zhang[#], Wan Hua, and Peter ten Dijke*

Oncode Institute, Leiden University Medical Center, Leiden, the Netherlands.

[#] These authors contributed equally.

*Correspondence: Peter ten Dijke

Department of Cell and Chemical Biology and Oncode Institute, Leiden University Medical Center, Einthovenweg 20, 2300 RC, Leiden, The Netherlands

Email: p.ten_dijke@lumc.nl; Telephone: +31 71526 9271; Fax: +31 71 526 8270

Introduction

Transforming growth factor (TGF)- β is a multifunctional secreted cytokine that exerts highly context dependent effects on many different cell types, including growth inhibition, extracellular matrix (ECM) production, apoptosis and differentiation^{1, 2}. TGF- β 1 is the prototype of a large family of evolutionarily conserved structurally and functionally related dimeric proteins that include TGF- β s, activins and bone morphogenetic proteins (BMPs). Signaling occurs via transmembrane serine/threonine kinase type I and type II receptors, that is T β RI and T β RRII, respectively³. TGF- β induces the formation of a complex of T β RI and T β RRII, upon which T β RRII phosphorylates T β RI, thereby transmitting the signal across the cell membrane. Inside the cells, activated T β RI phosphorylates specific down-stream effector molecules, among which are canonical SMAD and non-SMAD signaling components. SMADs can act as transcription factors and thus relay the signal from the membrane into the nucleus⁴. Each step of the signaling pathway is intricately regulated to fine tune the cellular responses of TGF- β ⁵.

Misregulation of TGF- β signaling associates with many diseases, including cancer, fibrosis and cardiovascular diseases⁶⁻⁸. In this review, we focus on its dual role in cancer. Moreover, as TGF- β stimulates cancer cell invasion and metastasis, this pathway has been subject to therapeutic targeting by academic and industrial laboratories. We provide an update on the latest clinical developments of TGF- β targeting agents for the treatment of cancer^{9, 10}.

TGF- β Signaling

Ligands and Their Receptors

Shortly after the cDNA cloning of TGF- β 1 in 1985¹¹, the structurally and functionally related TGF- β 2 and TGF- β 3 were characterized¹². In this review, we indicate specific TGF- β isoforms when relevant, for example, when they have distinct functional properties; otherwise we refer to them as TGF- β . TGF- β is a conserved 12.5 kilodalton (kDa) polypeptide that forms a disulfide-linked dimer¹³. While predominantly present as homodimers, heterodimers between different TGF- β isoforms have been described¹⁴. Of note, TGF- β may exert diverse, sometimes even opposing, effects depending on cell types and development stages^{1, 2}. The three TGF- β isoforms are differentially expressed. TGF- β 1 is highly abundant in platelets and bone and is widely expressed and synthesized among diverse tissues. TGF- β is secreted in an inactive form in which the amino-terminal pro-peptide (also termed the latency-associated peptide) is non-covalently associated with the carboxy-terminal mature peptide¹⁵. Activation of TGF- β can be

mediated via specific proteases and cell surface-associated integrins that liberate the mature peptide, which can then bind to cell surface receptors¹⁶. This activation step is a pivotal control mechanism that regulates the local bioavailability of TGF- β .

Activated TGF- β initiates cellular responses by binding to cell surface single transmembrane T β RI and T β RII¹⁷. TGF- β induces the formation of a heterotetrameric complex containing two T β RIIs and two T β RIIs¹⁸. Initially, TGF- β 1 and TGF- β 3 (but not TGF- β 2) bind to T β RII, and thereafter, T β RI is recruited. TGF- β type III coreceptor (also termed betaglycan), which lacks intercellular enzymatic activity, can facilitate the interaction between T β RI and T β RII¹⁹. In particular, TGF- β 2 requires T β RIII for efficient binding to signaling receptors²⁰. Upon the ligand-induced T β RI/T β RII complex formation, T β RI is phosphorylated by T β RII on specific serine and threonine residues in the glycine/serine-rich (GS) domain. The extracellular ligand signal is thereby transduced across the membrane, and the activated T β R complex is ready to initiate intracellular responses by phosphorylating intracellular effector proteins²¹⁻²³ (**Fig. 1**).

TGF- β /SMAD and Non-SMAD Signaling

With the help of genetic approaches in worms and fruit flies, Sma- and Mad-related proteins, termed SMAD proteins, were identified in vertebrates as unique and pivotal intracellular effectors of TGF- β ²⁴. SMADs are classified into three groups: the receptor-regulated SMADs (R-SMADs), the common SMADs (Co-SMADs) and the inhibitory SMADs (I-SMADs)²⁵⁻²⁷. R- and Co-SMADs share two conserved domains, i.e. N-terminal Mad Homology 1 (MH1) and C-terminal Mad Homology 2 (MH2) domain. Both domains are separated by a proline-rich linker region. There is also an MH2 domain in I-SMADs.

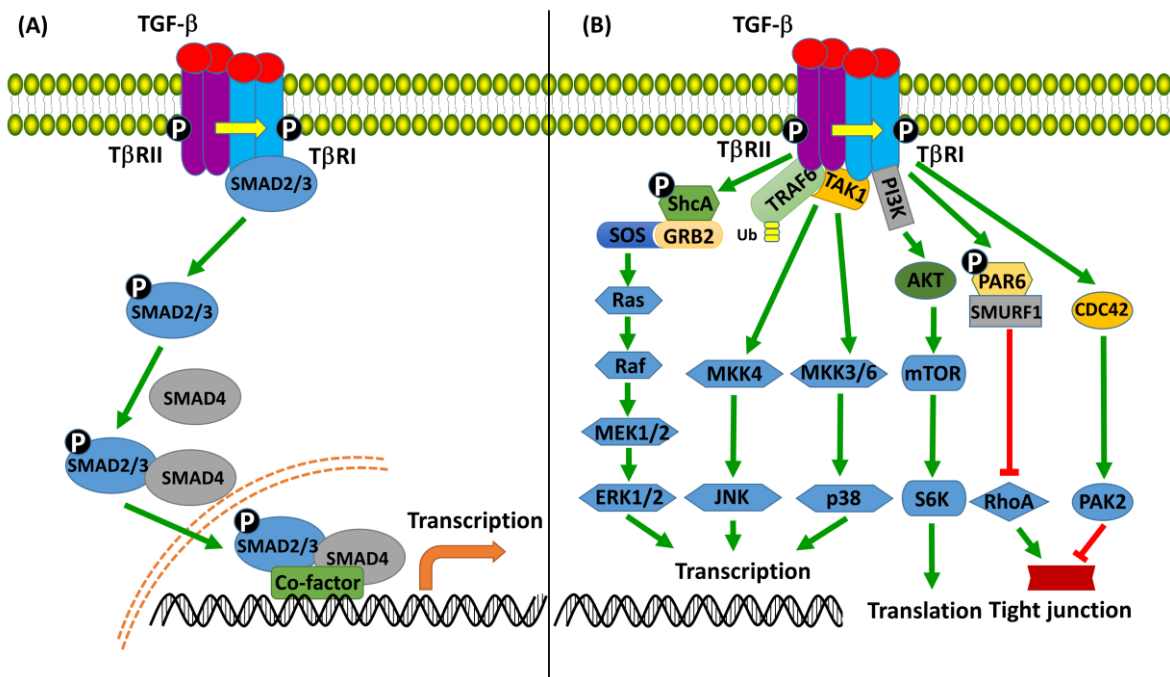


Fig. 1 TGF- β /SMAD and non-SMAD signaling. (A) In the SMAD-dependent pathway, binding of active TGF- β induces the assembly of T β RI and T β RII into a complex in which T β RI is phosphorylated by the T β RII kinase. Activated T β RI subsequently signals by recruiting and phosphorylating SMAD2/3, which form heteromeric complexes with SMAD4. The SMAD complexes then translocate into the nucleus and regulate target gene transcription by cooperating with other cofactors. (B) In the non-SMAD signaling pathways, TGF- β receptors activate other pathways including MAPKs (such as ERKs, p38 and JNK) and PI3K-AKT signaling to regulate transcriptional and translational events and modulate the Rho-like GTPase activity for tight junction dissolution. Abbreviations: ERK, extracellular regulated kinase; GRB2, growth factor receptor-bound protein 2; mTOR, mammalian target of rapamycin; PI3K, phosphatidylinositol-3-kinase; S6K, S6 kinase; SMURF,

SMAD ubiquitin regulatory factor; SOS, son of sevenless; TAK1, TGF- β activated kinase; T β R, TGF- β receptor; TGF- β , transforming growth factor- β ; TRAF, TNF associated factor; Ub, ubiquitin.

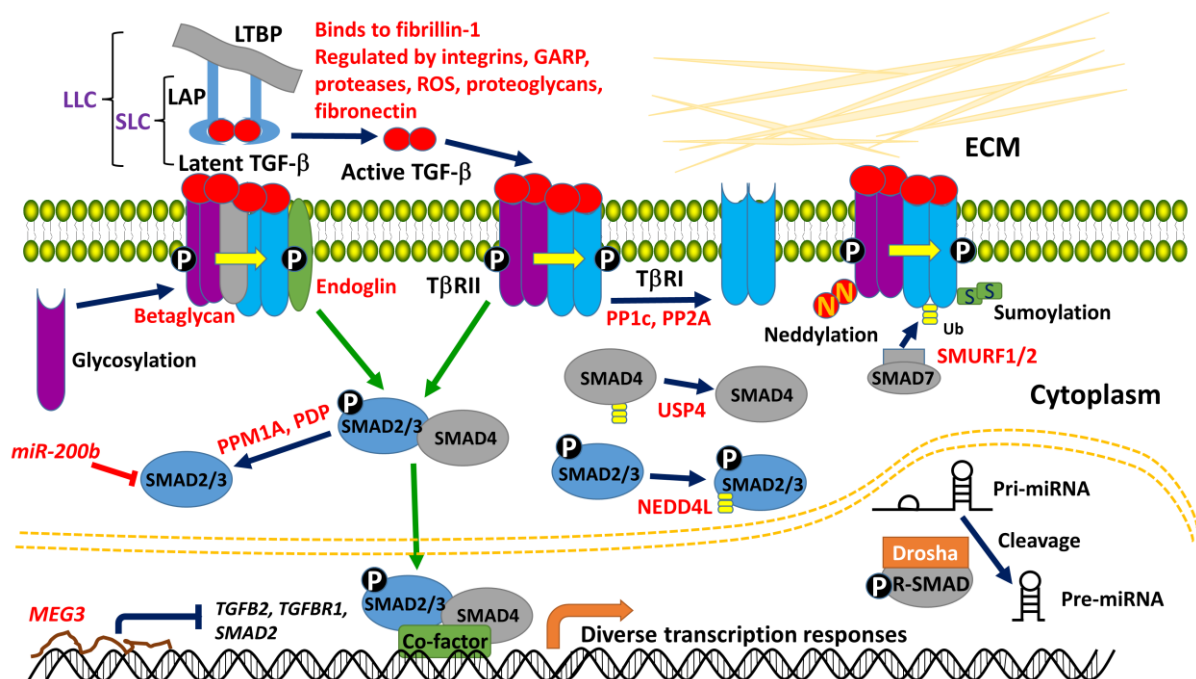


Fig. 2 Regulation of TGF- β /SMAD signaling. Fibrillin-1, proteases, ROS, GARP, integrin-mediated contractile forces and stromal-derived factors modulate the bioavailability of TGF- β ligands and accessibility to its receptors. At the cell membrane level, the activity of T β Rs is modified by glycosylation, phosphorylation, ubiquitylation, deubiquitylation, sumoylation and neddylation, as well as the interactions with coreceptors and other accessory proteins. At the cytoplasmic level, SMAD proteins are under tight control by phosphatases, ubiquitylating enzymes, deubiquitylating enzymes and microRNAs (miRNAs). In the nucleus, the SMAD complex affects different transcriptional responses in combination with diverse cofactors. SMAD proteins are also required for the maturation process of miRNAs. Moreover, modulators such as long non-coding RNAs (lncRNAs) can regulate TGF- β pathway components at the transcription level.

Upon activation, T β RI recruits and phosphorylates R-SMAD family members, SMAD2 and SMAD3, at two serine residues in their C-terminal regions. Activated SMAD2/3 form heteromeric complexes with SMAD4 which subsequently translocate into the nucleus. Activated SMAD2/3/4 complexes can form transcriptional complexes in conjunction with a large variety of DNA binding cofactors and thereby gain high affinity and specificity to DNA. The intrinsic binding activity of SMAD3 and SMAD4 (via their MH1 domain) is weak, and their direct binding ability to DNA is lacking in the predominantly expressed splice variant of SMAD2. These SMAD-containing transcription factor complexes interact with coactivators, corepressors and chromatin remodeling factors to regulate the transcription of target genes in a cell type-dependent manner^{22, 25, 28} (**Fig. 1A**).

In addition to the canonical SMAD-dependent pathway, non-SMAD signaling pathways can be initiated by activated TGF- β receptor complexes in specific cell types (**Fig. 1B**). These pathways can also modulate the SMAD pathway²⁹. Via phosphorylation or direct interaction with signaling modules, TGF- β receptors can activate pathways such as the mitogen-activated protein kinase (MAPK) signaling cascade, which includes extracellular signal-regulated kinases (ERKs), p38 and c-Jun amino terminal kinase (JNK), phosphatidylinositol-3 kinase (PI3K)-AKT signaling and Rho-like GTPase activity²⁹⁻³². T β RII is phosphorylated by non-receptor tyrosine kinase Src on Tyr284, which acts as a docking site for growth factor receptor-bound protein 2 (GRB2) and Src homology domain 2 containing (Shc), leading to the activation of ERK MAPK pathway³³. Moreover, Shc is reported to be directly phosphorylated by T β RI,

which provides a docking site for GRB2 that interacts with the exchange factor SOS to activate the pro-oncogenic Ras-Raf-MEK1/2-ERK1/2 signaling³¹. Phosphorylated ERK1/2 translocate into the nucleus and regulate gene transcription by phosphorylating target transcription factors³⁴. TGF- β activated kinase 1 (TAK1), a MAP kinase kinase kinase (MAPKKK) that is recruited to the TGF- β receptor complex by polyubiquitylated TRAF6, phosphorylates specific MAP kinase kinases (MKKs), leading to the phosphorylation of JNK and p38³⁵. In addition, TGF- β stimulation triggers the interaction between T β RI and the PI3K subunit p85, leading to AKT phosphorylation and the activation of downstream effectors (e.g., mTOR, P70S6K and 4EBP1)^{36,37}. PAR6 can also be phosphorylated by T β RI and recruit SMURF1 to degrade RhoA, which regulates cell–cell interactions via tight junctions³⁸. CDC42, another GTPase, can be recruited to the TGF- β receptor complex and mediate the activation of p21-activated kinase 2 (PAK2), which stimulates tight junction disassociation^{39,40} (**Fig. 1B**).

Regulation of TGF- β /SMAD Signaling

As a pivotal cytokine in cell homeostasis, TGF- β signaling activity is under precise control, from ligand bioavailability to receptor and SMAD activation (**Fig. 2**). After synthesis and intracellular furin-mediated cleavage of the precursor protein (removal of the signal peptide), the bioactive growth-factor domain (mature TGF- β) and prodomain, also termed the latency-associated peptide (LAP), are secreted in a small latent complex (SLC) form. Binding of TGF- β ligand to its receptors is prevented by LAP. The large latent complex (LLC), a more commonly deposited complex, contains the SLC and the latent TGF binding protein (LTBP)⁴¹⁻⁴⁴. LLC is bound to elastic microfibrils via the binding of LTBP to the extracellular protein fibrillin-1⁴⁵. Stromal-derived molecules including proteases and reactive oxygen species (ROS) substantially contribute to the increase of active TGF- β levels by interacting with the latent TGF- β complex^{43,46-48}. Moreover, glycoprotein-A repetitions predominant protein (GARP) functions as a critical docking receptor on regulatory T cells to concentrate and activate latent TGF- β on the cell surface^{49,50}. In addition, contractile forces exerted by the integrins across the LLC play a vital role in the release of mature TGF- β ^{42,51-53}. Fibronectin deposited in the ECM prior to LLC formation impairs TGF- β 1 bioactivity by interacting with LTBP⁵⁴. Decorin, a member of the proteoglycan family, also exerts a suppressive role in TGF- β activity via binding to all isoforms of soluble TGF- β ⁵⁵.

Apart from the ECM level, TGF- β responsiveness is tightly controlled at the cell membrane. Glycosylation of the extracellular domain of T β RII inhibits its transportation to the cell membrane and lowers its TGF- β binding affinity^{56,57}. E3 ubiquitin ligases such as SMAD-specific E3 ubiquitin protein ligase 1/2 (SMURF1/2) cooperate with inhibitory SMAD7 to regulate the availability of T β RI receptor on the cell surface by polyubiquitylation and proteasomal degradation^{58,59}. In contrast, deubiquitinating enzymes ubiquitin-specific protease (USP) 4, 11 and 15 remove the polyubiquitin chains from T β RI⁶⁰. Moreover, two phosphatases, i.e. protein phosphatase (PP)1c and PP2A, impair receptor activation by targeting T β RI for dephosphorylation^{61,62}. Akin to ubiquitylation, sumoylation and neddylation have also been implicated to regulate TGF- β receptor stability. The interaction between TGF- β receptors and the coreceptors located in the cell membrane is another determinant for the signaling strength^{21,63}. The coreceptor betaglycan stabilizes the receptor complex between T β RI and T β RII and propagates signaling transduction initiated by TGF- β 2⁶⁴. Endoglin, another accessory protein structurally related to betaglycan, inhibits TGF- β /ALK5-mediated SMAD2/3 signaling but promotes TGF- β /ALK1-induced SMAD1/5/8 signaling in endothelial cells^{65,66}.

At the cytoplasmic level, phosphorylated SMAD proteins can be deactivated by phosphatases such as PPM1A and PDP, leading to signal termination⁶⁷⁻⁶⁹. Similar to the TGF- β receptors,

SMAD2/3 are destabilized by multiple ubiquitylating enzymes such as SMURF1/2 and NEDD4L^{70, 71}. Conversely, USP4 promotes SMAD4 activity by removing the suppressive monoubiquitination triggered by SMURF2⁷². MicroRNAs (MiRNAs) also inhibit the expression of various signaling components⁷³. *MiR-200b*, a miRNA whose expression is downregulated by TGF- β 1, attenuates TGF- β signaling by targeting *SMAD2* mRNA at the post-transcriptional level, thereby forming a negative feedback loop⁷⁴.

Upon activation, the SMAD2/3/4 complex translocate into the nucleus and form a transcription complex with other cofactors. In combination with different sequence-specific transcription factors, the SMAD complex generate various transcriptional responses in a context and cell type-dependent manner^{1, 75-77}. In addition, activated SMAD proteins participate in the maturation of miRNAs by recruiting the RNA helicase p68 (DDX5) to the Drosha complex⁷⁸. *MEG3*, an intranuclear long noncoding RNA (lncRNA), can bind to the distal regulatory elements of genes encoding for TGF- β signaling components, including *TGFB2*, *TGFBR1* and *SMAD*, to inhibit their transcription⁷⁹.

TGF- β as a Tumor Suppressor

TGF- β -Induced Growth Inhibition

TGF- β induces growth inhibition (**Fig. 3**) and apoptosis (**Fig. 4**) in normal epithelial (and certain premalignant) cells; these properties are associated with its function as a tumor suppressor⁸⁰. The molecular mechanisms by which TGF- β elicits these processes involve multiple intracellular pathways⁸¹⁻⁸³.

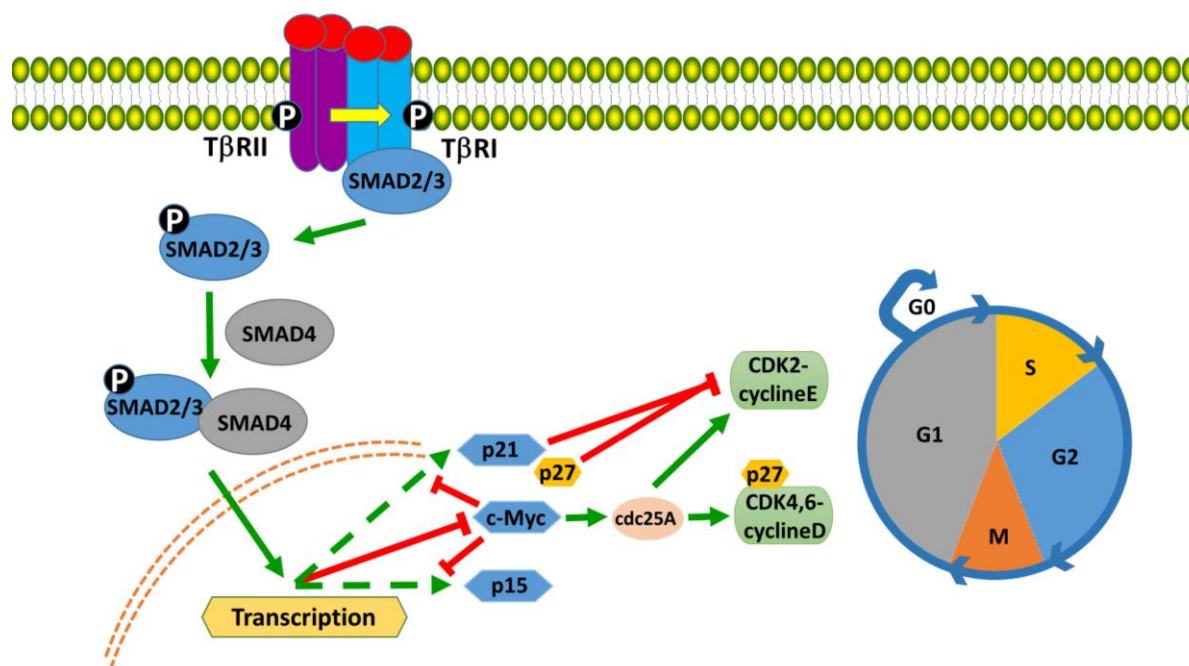


Fig. 3 Gene regulation in TGF- β -induced cell cycle arrest. TGF- β receptor activation leads to SMAD2/3 phosphorylation. Phosphorylated SMAD2/3 bind to SMAD4, and the SMAD2/3/4 complex translocate into the nucleus to modulate gene transcription. *C-Myc* and *cdc25A* gene expression is repressed, while *p15INK4b* and *p21CIP1/WAF1* gene expression is induced by TGF- β , leading to the cell cycle arrest into the G1 phase. Abbreviations: CDK, cyclin dependent protein kinase; T β R, TGF- β receptor; TGF- β , transforming growth factor- β .

Numerous studies support the notion that TGF- β inhibits cell proliferation by arresting cells into the G1 phase of the cell cycle (**Fig. 3**). SMAD-containing protein and transcriptional coactivator complexes can activate the transcription of two major cell cycle inhibitors, CDK

inhibitors (CKIs), *p15* and *p21*^{84, 85}. In keratinocytes, TGF- β /SMAD signaling induces the expression of cyclin-dependent kinase inhibitors *p15*^{INK4b} and *p21*^{CIP/WAF1}, which inhibit the CDK4/6-cyclinD complex⁸⁶. These cyclin-dependent kinase inhibitors suppress the CDK activities associated with the G1 to S phase progression, prevent cyclin-dependent kinases-mediated Rb phosphorylation, and arrest cells in the G1 phase⁸⁷. The activated SMAD proteins target the promoters of *c-Myc* and *CDK* genes and repress their transcription in cooperation with nuclear corepressors⁸⁸. TGF- β receptor-initiated non-SMAD signaling can also exert an anti-proliferative effect on some cell types⁸⁹.

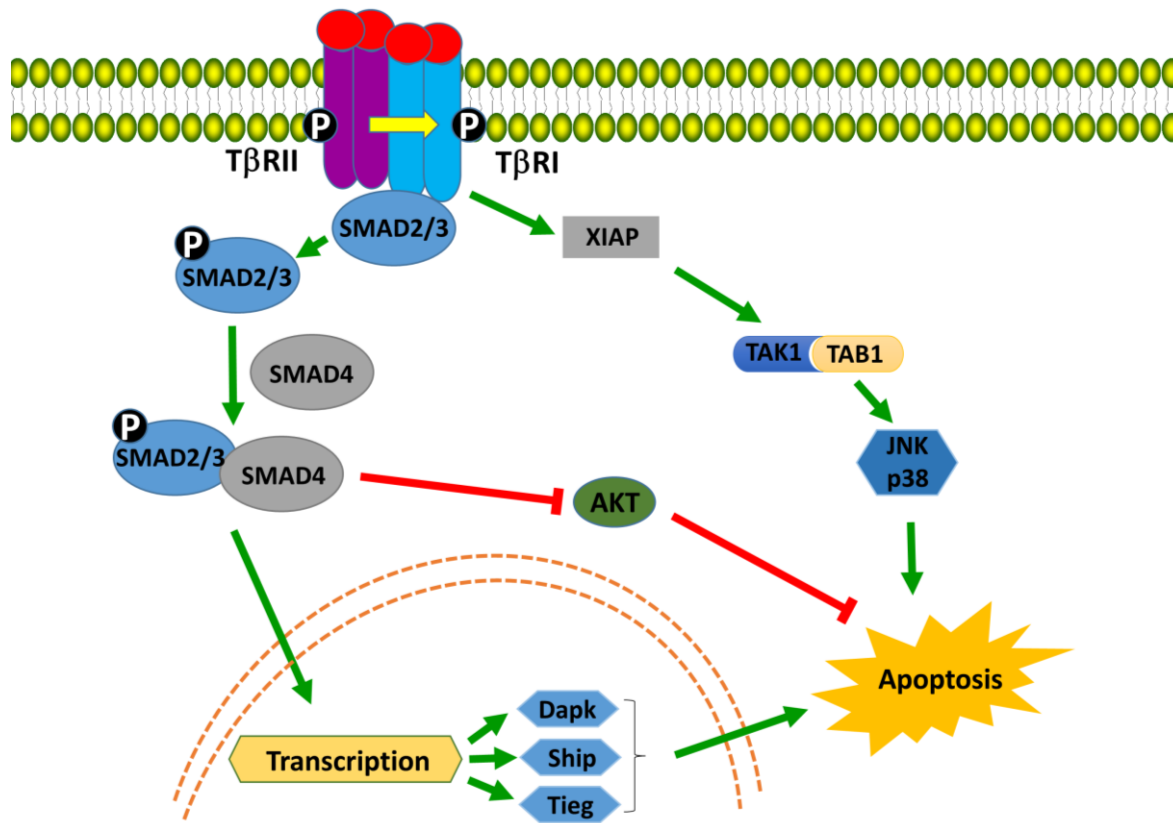


Fig. 4 TGF- β -induced cell apoptosis. TGF- β promotes the activation of SMADs and the expression of pro-apoptotic genes such as *Dapk*, *Ship* and *Tieg*. SMADs also bind and inactivate the survival kinase AKT, thereby inducing apoptosis. TGF- β -induced activation of the JNK and p38 pathways can also result in apoptosis. TGF- β can also induce, via the adaptor XIAP, the activation of the TAK1-TAB complex, leading to JNK or p38 activation, both of which can lead to apoptosis. Abbreviations: Dapk, death associated protein kinase; Ship, SH2-containing inositol phosphatase; TAB1, TAK1 binding protein; TAK1, TGF- β activating kinase; TGF- β , transforming growth factor- β ; Tieg, TGF- β -inducible early-response gene; XIAP, X chromosome-linked inhibitor of apoptosis.

TGF- β -Induced Apoptosis

TGF- β can induce cell apoptosis in normal epithelial (and some premalignant) cells (Fig. 4). Several apoptotic regulators have been implicated as downstream targets of TGF- β signaling, often in a cell- or tissue-specific manner⁹⁰. Induction of the pro-apoptotic genes such as *Ship* and *Tieg* have been shown in TGF- β -induced apoptosis⁹¹. In liver cancer cells, the Daxx adaptor protein couples TGF- β signaling to the cell death machinery through its interaction with T β RII⁹². In liver cancer cells, TGF- β can induce the expression of the death-associated protein kinase *DAPK*, which promotes cell death⁹³. In addition, TGF- β -induced activation of TGF- β -activated kinase-1 (TAK-1), a protein of the MAPKKK family that activates p38 and JNK signaling, is involved in TGF- β -induced apoptosis⁹⁴. TGF- β can also induce apoptosis

through repressing the phosphoinositide 3-kinase/AKT/survivin pathway in colon cancer cells⁹⁵ (Fig. 4).

Mutation in TGF- β Signaling Components in Cancer

Analysis from clinical tumor samples reveals that TGF- β -mediated signaling is indeed strongly implicated in the regulation of cancer⁹⁶. Recent studies have shown that in various human tumor types, components of the TGF- β signaling pathway, namely, *TGFBR2*, *TGFBR1*, *SMAD2*, *SMAD3* and *SMAD4*, are commonly inactivated through mutation^{81, 97}. Multiple genetic alterations in genes encoding central components in TGF- β signaling pathway are found in human cancers, in particular in pancreatic, esophagus, colorectal and head and neck cancer⁹⁸ (Fig. 5). Indeed, *TGFBR2*-inactivating mutations in its poly A gene tract are frequently found in cancers associated with microsatellite instability (MSI)⁹⁹. SMAD point mutations associated with cancer are loss-of-function mutations that either target functional elements or affect the overall stability of the protein. Studies in cultured cells have shown that these inactivating mutations mediate an escape from TGF- β -induced growth arrest and apoptosis.

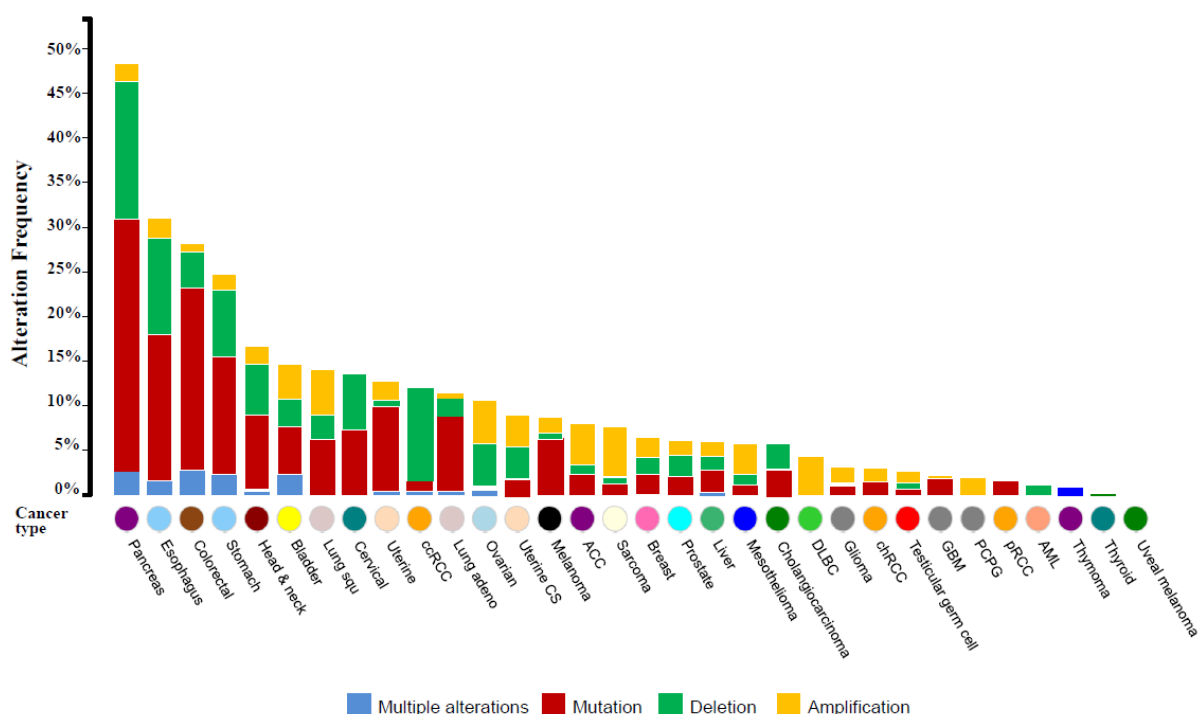


Fig. 5 Frequency of genetic alterations in *TGFBR1*, *TGFBR2*, *SMAD2*, *SMAD3* and *SMAD4* by cancer type. The graph displays the frequency of genetic alterations (point mutations, deletions, amplifications, or multiple alterations) in *TGFBR1*, *TGFBR2*, *SMAD2*, *SMAD3* and *SMAD4* in different types of cancer. Data were derived from TCGA datasets (The Cancer Genome Atlas, cancergenome.nih.gov/) at the time of this writing. Analysis was done using cBioPortal (www.cbioportal.org/).

In addition to the known mutations in the TGF- β receptors and SMAD pathway, other types of (epi)genetic alterations may also affect TGF- β signaling and tumor formation⁸⁹. For example, oncogenic activation of the Ras-Raf-MAPK pathway and c-Jun NH₂-terminal kinase in hepatocellular carcinoma has been reported to induce phosphorylation of the SMAD3 linker domain by MAPK, further preventing C-terminal phosphorylation of SMAD by the T β RI kinase domain and thereby inhibiting the TGF- β cytostatic effects¹⁰⁰.

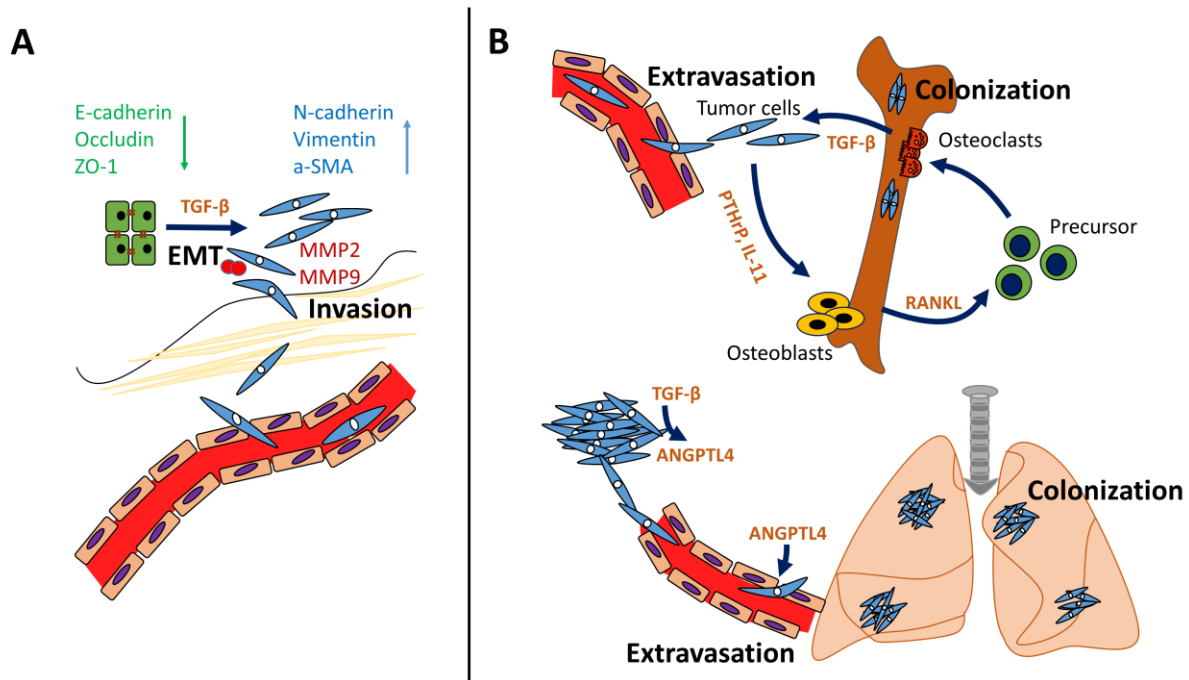


Fig. 6 TGF- β -induced EMT, invasion and metastasis. (A) TGF- β induces EMT by decreasing the expression of epithelial makers (in green) and increasing the expression of mesenchymal markers (in blue). TGF- β also promotes the secretion of MMP2 and MMP9, thereby conferring tumor cells highly invasive abilities. (B) Bone-derived TGF- β increases the secretion of PTHrP, which activates osteoclast activity through interacting with RANKL, thereby promoting osteolytic metastasis. IL-11 and CTGF are also key effectors induced by TGF- β in this process. Osteolysis leads to more local TGF- β release, causing the formation of a positive feedback loop. Moreover, TGF- β -induced ANGPTL4 plays a vital role in disrupting the junctions between pulmonary endothelial cells and contributes to lung metastasis formation. Abbreviations: ANGPTL4, angiopoietin-like 4; EMT, epithelial to mesenchymal transition; MMP, matrix metalloproteinase; SMA, smooth muscle actin; RANKL, receptor activator of nuclear factor κ B ligand; TGF- β , transforming growth factor- β .

TGF- β as a Tumor Promoter

TGF- β -Induced EMT and Invasion

In the late stage of tumor progression, TGF- β switches from a tumor suppressor to a tumor promoter by inducing EMT, tumor invasion, distant dissemination, angiogenesis and immune evasion¹⁰¹⁻¹⁰⁴. During EMT, tumor cells switch from an epithelial phenotype to a mesenchymal phenotype and gain highly migratory and invasive abilities. Moreover, they acquire cancer stem cell (CSC) properties and become more resistant to detachment-induced apoptosis¹⁰⁵. During EMT, epithelial cells downregulate the expression of genes encoding epithelial markers, such as E-cadherin, Occludin and ZO-1, upregulate the expression of genes encoding mesenchymal markers, such as N-cadherin, Vimentin and α -smooth muscle actin (SMA), and dissolve the tight junctions. EMT greatly facilitate tumor cell invasion^{106, 107} (**Fig. 6A**). In response to TGF- β , the SMAD complex directly increases the expression of multiple EMT-transcription factors including *ZEB*, *TWIST* and *SNAIL* family members by binding to their promoters. In addition, in combination with *ZEB2* or *SNAIL*, the SMAD complex suppresses the transcription of genes encoding E-cadherin and Occludin, conferring the mesenchymal traits to cancer cells^{108, 109}. In addition, SMAD4 binding enhances the promoter activity of *miR-155*. *MiR-155* dissolves the tight junctions by targeting *RhoA* mRNA and downregulates *CDH1* (mRNA encoding for E-cadherin) expression by inhibiting the expression of transcriptional activator *CEBPB* (mRNA encoding for C/EBP β)^{110, 111}. *LncRNA-ATB*, a long non-coding RNA activated by TGF- β , serves as a sponge for the *miR-200* family members that restrain *ZEB1/2* protein expression, and thereby promotes EMT and hepatocellular carcinoma progression¹¹². In combination with the SMAD-dependent pathway, SMAD-independent pathways also

potentiate TGF- β -induced EMT^{29, 107, 113}. Activation of proto-oncogenes such as Ras and receptor tyrosine kinase pathways cooperate with TGF- β pathway to promote EMT^{114, 115}. By directly modulating the activity of AP1 transcription factors that can cooperate with SMADs or phosphorylate R-SMADs, the ERK, p38 and JNK MAPK pathways play a key role in TGF- β -induced EMT and tumor invasion¹¹⁶⁻¹¹⁹. In addition, PI3K/AKT signaling participates in TGF- β -triggered EMT by activating the mTOR and EMT-related transcription factors such as SNAIL and TWIST1^{37, 120-122}. Activation of the Rho family GTPases including RhoA, Rac1, and Cdc42 by TGF- β receptors contributes to cell–cell junction dissolution and cytoskeletal reorganization, which are important determinants for EMT^{38, 123, 124}.

Local invasion through the surrounding ECM and stromal cell layers is the first step of the invasion-metastasis cascade¹²⁵. Results from human cancer specimens suggest that coexpression of SMAD3/4 and SNAIL is correlated with the loss of E-cadherin and coxsackie and adenovirus receptor (CAR), a tight junction-associated cell adhesion molecule, at the invasive front¹⁰⁹. Apart from conferring EMT properties to cancer cells, TGF- β induces the expression and secretion of *matrix metalloproteinases 2/9* (MMP2/9) in tumor cells or/and stromal cells (e.g., myofibroblasts). These two proteinases promote ECM and collagen proteolysis, leading to the invasion of tumor cells into their stromal compartment^{126, 127} (**Fig. 6A**). In addition, TGF- β employs *miR-181b* to inhibit the protein level of TIMP3, an inhibitor of metalloprotease. The latter promotes MMP2/9 activities and the invasion of hepatocellular carcinoma cells¹²⁸.

TGF- β -Induced Metastasis to Bone, Lung and Other Organs

Cancer metastasis contributes to the death of most cancer patients¹²⁹. Bone metastasis is a common event in specific cancer types, including breast, lung and prostate cancers. The interaction between disseminated cancer cells and resident skeletal cells disrupts bone integrity, conferring a receptive microenvironment for the outgrowth of metastatic cancer cells^{130, 131}. Bone-derived TGF- β promotes SMAD-dependent pathway activation in cancer cells, which increases the expression and secretion of *parathyroid hormone-related protein* (PTHrP), a major osteoclastogenic factor. PTHrP potentiates osteoclast activity by interacting with receptor activator of nuclear factor κ B ligand (RANKL), thereby promoting bone metastasis¹³¹⁻¹³⁴ (**Fig. 6B**). By employing *in vivo* selection of highly metastatic cell lines and functional imaging, Kang *et al.* identified a bone metastasis gene signature that includes *C-X-C motif chemokine receptor 4* (CXCR4), *interleukin 11* (IL-11) and *connective tissue growth factor* (CTGF), which contribute to metastasis by directing the homing of breast cancer cells to bone, osteolysis and angiogenesis, respectively^{135, 136}. IL-11 and CTGF expression is induced by TGF- β . The degraded bone in turn secretes stored factors including TGF- β to form a positive feedback loop called “vicious cycle”¹³⁷.

TGF- β signaling also contributes to lung metastases formation. A TGF- β -induced gene expression signature in estrogen receptor (ER)-negative breast cancer cells was found to correlate with the potential to form lung metastases. Blockade of TGF- β signaling impairs the extravasation of ER-negative breast cancer cells in lung capillaries, while TGF- β pretreatment increases the metastatic abilities of tumor cells¹³⁸. TGF- β -induced *adipokine angiopoietin-like 4* (ANGPTL4) plays a vital role in the disruption of junctions between pulmonary endothelial cells (**Fig. 6B**). However, bone metastasis is not affected by TGF- β preincubation or ANGPTL4 knockdown, which can be explained by the microvasculature difference in these two organs⁹⁶.

TGF- β also participates in the metastatic growth of tumor cells in liver^{139, 140}. Upon extravasating into liver parenchyma, TGF- β released by colorectal cancer cells promotes the

transformation of surrounding hepatic satellite cells (HSCs) into myofibroblasts. Tumor-associated myofibroblasts in turn increase the expression of *C-X-C motif chemokine ligand 12* (*CXCL12*) and *hepatic growth factor* (*HGF*), which trigger the metastatic growth of cancer cells¹⁴¹.

Stimulation of Angiogenesis and Immune Evasion by TGF- β

Angiogenesis is indispensable for solid tumors larger than 2–3 mm³ to obtain oxygen and nutrients, remove waste products and spread through the circulatory system¹⁴². An elevated level of TGF- β in plasma correlates with an increase of tumor angiogenesis and poor clinical outcomes in many cancer types¹⁴³⁻¹⁴⁶. TGF- β can directly activate endothelial cells by promoting TGF- β /ALK1 signaling¹⁴⁷. The coreceptor endoglin, which is highly expressed in activated endothelial cells, can potentiate this signaling response¹⁴⁸. Moreover, in the tumor niche with low oxygen, hypoxia and TGF- β signaling can cooperate to initiate an angiogenic program in cancer cells. Mechanistically, hypoxia-induced HIF-1, in cooperation with SMAD3, enhances the transcription of *vascular endothelial growth factor* (*VEGF*), which is of importance in capillary formation and endothelial cell migration, thereby promoting tumor angiogenesis^{149, 150}.

Table 1 Overview of clinical trials with TGF- β targeting agents

Stage	Drug	Type	Target	Disease	Clinical trial identifier
Phase 1	AP 12009	AON	TGF- β 2	Multiple cancers	NCT00844064
	TAG Vaccine	Vaccine therapy	TGF- β 2 & immune	Carcinoma/Advanced Metastatic	NCT00684294
	TEW-7197	Kinase inhibitor	T β RI	Advanced Stage Solid Tumors	NCT02160106
	NIS793	Antibody	TGF- β	Multiple cancers	NCT02947165
	Paclitaxel/Carboplatin + Galunisertib	Kinase inhibitor	T β RI	Carcinosarcoma, Ovarian	NCT03206177
	TEW-7197	Kinase inhibitor	T β RI	Multiple Myeloma	NCT03143985
	LY2157299	Kinase inhibitor	T β RI	Multiple myeloma	NCT00689507
	LY573636	Kinase inhibitor	T β RI	Hematopoietic malignancies	NCT00718159
	Fresolimumab (CG1008)	Antibody	TGF- β 2	Multiple cancers	NCT00356460
Phase 1/Phase 2	Fresolimumab (CG1008)	Antibody	TGF- β 2	Multiple cancers	NCT02581787
	LY2157299	Kinase inhibitor	T β RI	Multiple cancers	NCT02423343
	LY2157299	Kinase inhibitor	T β RI	Glioma	NCT01220271
Phase 2	LY2157299	Kinase inhibitor	T β RI	Prostate Cancer	NCT02452008
	Lucanix (belagen-pumatucl)	Vaccine therapy	TGF- β 2 & immune	Multiple cancers	NCT01058785
	PF03446962	Antibody	ALK1	Transitional Cell Carcinoma of Bladder	NCT01620970
	Fresolimumab (CG1008)	Antibody	TGF- β 2	Primary Brain Tumors	NCT01472731
	LY2157299	Kinase inhibitor	T β RI	Metastatic Breast Cancer	NCT02538471
	LY2157299	Kinase inhibitor	T β RI	Rectal Adenocarcinoma	NCT02688712
	AP 12009	AON	TGF- β 2	Glioblastoma/Anaplastic Astrocytoma	NCT00431561
	Fresolimumab (CG1008)	Antibody	TGF- β 2	Pleural Malignant Mesothelioma	NCT01112293
	Fresolimumab (CG1008)	Antibody	TGF- β 2	Renal Cell Carcinoma	NCT00923169
	LY2157299	Kinase inhibitor	T β RI	Hepatocellular Carcinoma	NCT02178358
Fresolimumab(CG1008)	Antibody	TGF- β 2	Metastatic Breast Cancer	NCT01401062	

	LY2157299	Kinase inhibitor	TβRI	Hepatocellular carcinoma	NCT01246986
	LY573636	Kinase inhibitor	TβRI	Melanoma	NCT00383292
	LY573636	Kinase inhibitor	TβRI	Non-small cell lung carcinoma	NCT00363766
	Fresolimumab(CG1008)	Antibody	TGF-β2 & immune response	Kidney cancer	NCT00899444
Phase 3	AP 12009	AON	TGF-β2	Anaplastic Astrocytoma/Glioblastoma	NCT00761280
	Lucanix	Vaccine therapy	TGF-β2 & immune response	Non-small Cell Lung Cancer	NCT00676507
Phase 4	Vitamin D3		TGF-β1	Multiple cancers	NCT02460380
Preclinical	LY2109761	Kinase inhibitor	TβRI/TβRII	Pancreatic cancer	
	SD208	Kinase inhibitor	TβRI	Melanoma	
	SM16	Kinase inhibitor	TβRI	Multiple cancers	
	TβRII Antibody	Antibody	TβRII	Multiple cancers	
	sTβRII (Fc)	Ligand Trap	TβRII	Multiple cancers	
	sBetaglycan	Ligand Trap	Betaglycan	Multiple cancers	
	1D11	Antibody	TGF-β1/2/3	Multiple cancers	
2G7	Antibody	TGF-β1/2/3	Multiple cancers		

Data from www.clinicaltrials.gov

Silencing *SMAD2* (in contrast to *SMAD3* depletion) in breast cancer MDA-MB-231 cells enhances TGF-β-induced VEGF secretion *in vitro* and promotes the formation of bone metastases *in vivo*¹⁵¹. TGF-β also enhances the transcription of *CTGF*, another key angiogenic factor, in breast cancer cells with high bone metastatic potential¹³⁵.

In addition to supporting EMT, invasion, metastasis and tumor angiogenesis, TGF-β also contributes to tumor progression by stimulating tumor evasion from immune surveillance. CD8⁺ cytotoxic T cells are a cell population that can induce cancer cell apoptosis. TGF-β represses the transcription of *granzyme*, *perforin* and *interferon-γ* through SMAD and ATF1 in CD8⁺ T cells, thereby inhibiting the cytotoxic activity of CD8⁺ cytotoxic T cells^{152, 153}. TGF-β can also induce the differentiation of regulatory T-cells (Tregs), which suppress the proliferation and activation of CD8⁺ cytotoxic T cells, resulting in immunosuppression and a decrease in immunosurveillance¹⁵⁴⁻¹⁵⁶. The activation of natural killer (NK) cells, another cytotoxic cell type, is attenuated by TGF-β-induced downregulation of *IL-15* and *NKG2D*, an activating receptor of NK cells^{157, 158}. In addition, TGF-β-triggered *miR-183* expression represses DAP12 protein expression, leading to the destabilization of the NK receptor and inhibition of cytotoxicity¹⁵⁹. In addition, TGF-β is a driver of the tumor-suppressive M1 macrophage phenotype transition into the tumor-promoting M2 phenotype, thereby promoting the production of tumor-promoting factors and inhibiting the activity of T cells^{160, 161}.

Targeting TGF-β Signaling in Cancer

Due to the strong pro-oncogenic effects of TGF-β, inhibitory agents targeting TGF-β have been developed, including antisense oligonucleotides (AONs), small molecule receptor kinase inhibitors and neutralizing antibodies. The mechanisms of these inhibitors involve the inhibition of TGF-β and receptor expression, the interference of receptor kinase signaling, and the blockade of TGF-β ligand and receptor binding (**Fig. 7, Table 1**). These agents have been tested in preclinical and clinical stages. While inhibiting tumor progression by blocking TGF-β signaling is a promising approach, the biphasic action of TGF-β in cancer progression and its multifunctionality make it a challenging target.

Antisense Oligonucleotides

The binding of ligands to receptors is the first step of TGF- β signaling activation; AONs have been developed to degrade *TGFB* mRNA¹⁶² (**Fig. 7**). The antisense RNA drugs AP12009 and AP11014 targeting *TGFB2* and *TGFB1*, respectively, have been used in (pre)clinical cancer treatment studies. AP12009 has been reported to inhibit neovascularization and tumor invasion and has been used to treat high-grade glioma and anaplastic astrocytoma patients¹⁶³⁻¹⁶⁵. In addition, AP11014 has been reported to display an anti-tumor effect in animal models of colon cancer, prostate cancer and lung cancer and is being studied in preclinical research^{166, 167}.

TGF- β Receptor Kinase Inhibitors

Small ATP-mimetic compounds have been synthesized to selectively inhibit T β RI (and T β RII) kinase activity (**Fig. 7**). These compounds have been tested in preclinical and clinical studies of multiple cancer types. Systemic administration of the T β RI kinase inhibitor SD208 can increase the median survival of mice with malignant glioma inoculation¹⁶⁸ and reduce tumor metastasis in pancreatic and breast cancer^{169, 170}. LY2157299 is the first T β RI kinase inhibitor that has been reported to inhibit primary tumor growth in breast and lung cancer cell lines^{171, 172}. To optimize the applicability of LY2157299 to cancer therapy, a first-in-human dose evaluation found that LY2157299 administration at 300 mg per day is safe¹⁷³. Another kinase inhibitor, LY2109761, inhibits both the activity of T β RI and T β RII. A large number of studies have indicated that LY2109761 exhibits great potential in the prevention of cancer metastasis in multiple cancer types including colon¹⁷⁴ and pancreatic cancer¹⁷⁵, glioblastoma¹⁷⁶ and ovarian cancer^{177, 178}. TEW-7197 is an orally administered small molecule that targets T β RI kinase activity. It stimulates apoptosis and suppresses TGF- β -induced activation of SMAD2/3 in human and murine myeloma cells *in vitro*, leading to the inhibition of myeloma cell growth and viability¹⁷⁹. While the preclinical results of these studies are promising, the clinical translation has been difficult. On-target side effects on the cardiovascular system have halted clinical advancement. By using an intermittent dosing strategy, these adverse side effects may be overcome^{172, 180}.

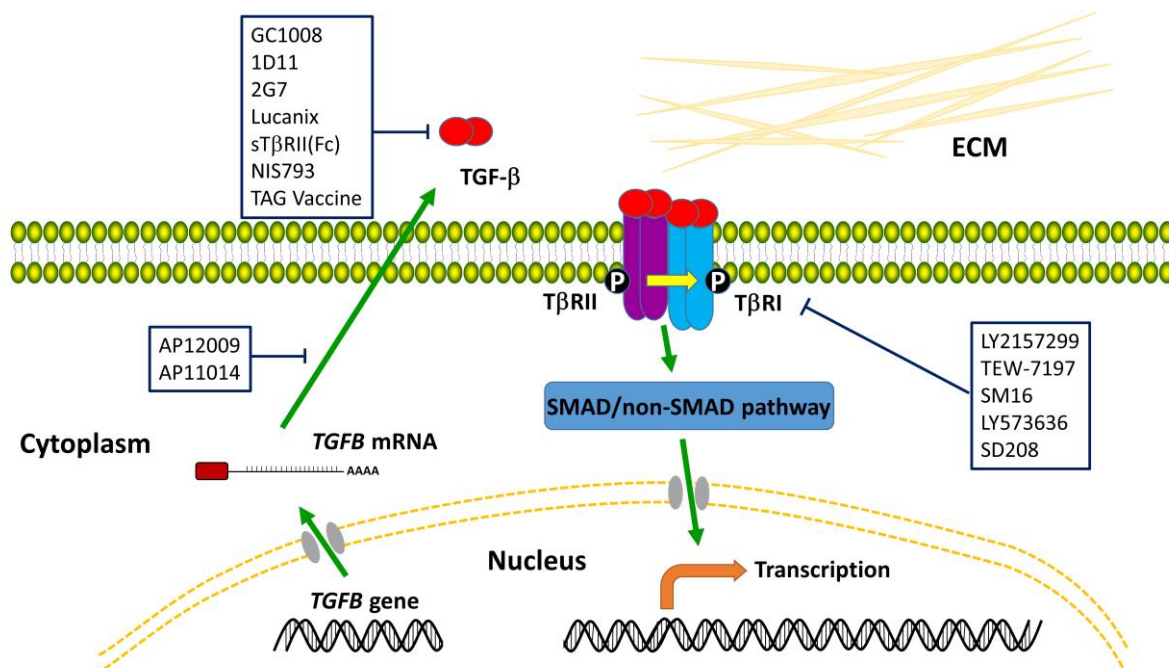


Fig. 7 Targeting TGF- β in cancer. TGF- β signaling has an important effect on tumor progression and provides a new approach for tumor targeting therapy. Many inhibitors of the TGF- β pathway (including kinase inhibitors, AONs, and antibodies) have already been applied in preclinical and clinical trials (see **Table 1**). Abbreviations: ECM, extracellular matrix; T β R, TGF- β receptor; TGF- β , transforming growth factor- β .

Antibodies Against TGF- β Ligands and Extracellular Domains of TGF- β Receptors

1D11, an antibody that recognizes all three TGF- β isoforms, interferes with TGF- β and TGF- β receptor binding and, thus, neutralizes TGF- β activity (**Fig. 7**). This antibody has been used in (pre)clinical studies. 1D11 significantly increases NK cell and nuclear T cell invasion, as well as *NKG2D* expression and cytotoxic perforin and granzyme B release in breast cancer cells, thereby enhancing the anti-tumor effect of CD8⁺ T cells and NK cells¹⁸¹. Additionally, 1D11 has also been found to suppress bone metastasis in prostate cancer¹⁸². Like 1D11, 2G7 also inhibits MDA-MB-231 cell invasion. Additionally, the combination of dendritic cell (DC)-based vaccines and 2G7 potently inhibits the development of the established murine mammary tumors^{183, 184}. A clinical trial of another monoclonal antibody, fresolimumab (CG1008), which inhibits all three TGF- β isoforms, demonstrated its safety and efficacy in suppressing metastatic melanoma and renal cell carcinoma¹⁸⁵. Fresolimumab may help to stabilize the condition of patients during malignant pleural mesothelioma therapy. Importantly, adverse effects, such as skin toxicity (including formation of cutaneous squamous-cell carcinomas and basal cell carcinoma), have been reported in cancer patients after fresolimumab treatment¹⁸⁵.

Similar to neutralizing antibodies, soluble T β RII and T β RIII ligand traps are also used to block TGF- β signaling. These molecules are expressed in the extracellular domain of the receptor, which prevent ligands from binding to TGF- β receptors¹⁰³. The ligand trap T β RII:Fc (a fusion of the extracellular TGF- β -binding domain of T β RII with IgG1 Fc domain) shows anti-tumor effects on multiple cancers, including inhibition of mesothelioma growth and suppression of breast cancer cell viability and migration^{186, 187}. The expression of soluble T β RIII (sBetaglycan) effectively suppresses tumor growth in MDA-MB-231 xenograft-bearing athymic nude mice¹⁸⁸ and inhibits glioma and non-small cell lung cancer progression in other mouse models^{189, 190}. Due to the risk of tumor development caused by the TGF- β soluble receptors^{191, 192}, these receptors have not yet entered the clinical research phase.

Targeting TGF- β signaling provides a new approach and opportunity for cancer therapy. Since TGF- β pathway is also involved in many normal biological functions, the exact mechanism of action in the patients and the adverse reactions caused by systemic inhibition of TGF- β are still not clear. A further understanding of the dual roles of TGF- β will be beneficial to the development of therapeutics specifically targeting TGF- β in tumor progression. Sole treatment with TGF- β -targeting agents will likely not be successful in curing cancer patients, and a combination of TGF- β targeting therapies with chemo- and radiotherapy or other forms of targeted therapy should be explored.

Concluding Remarks

TGF- β has a dual action in cancer by acting as a tumor suppressor in the early stages and a tumor promoter in the late phases of tumor progression. Cancer cells are insensitive to the cytostatic effects of TGF- β through the activation of proto-oncogenes and inactivation of tumor suppressor genes. The latter (epi)genetic changes also cooperate with TGF- β to mediate EMT, thereby facilitating invasion and metastasis. Moreover, TGF- β promotes tumorigenesis by stimulating immune evasion and promoting angiogenesis. The biphasic role in cancer and its multifunctional properties in the maintenance of tissue homeostasis make TGF- β a challenging pathway to target for treatment of cancer patients. A more detailed understanding of the mechanism of action in cancer patients, careful dosing and the selection of patients who will most benefit from the TGF- β targeting agents will be important for their clinical implementation.

References

1. Massagué, J. TGF β signalling in context. *Nature reviews Molecular cell biology* **13**, 616-630 (2012).
2. Ayyaz, A., Attisano, L. & Wrana, J.L. Recent advances in understanding contextual TGF β signaling. *Fl1000Research* **6**, 749 (2017).
3. Feng, X.H. & Derynck, R. Specificity and versatility in TGF- β signaling through SMADs. *Annual review of cell and developmental biology* **21**, 659-693 (2005).
4. Hill, C.S. Transcriptional control by the SMADs. *Cold Spring Harbor perspectives in biology* **8**, a022079 (2016).
5. Xu, P., Lin, X. & Feng, X.H. Posttranslational regulation of SMADs. *Cold Spring Harbor perspectives in biology* **8**, a022087 (2016).
6. Zhang, Y., Alexander, P.B. & Wang, X.F. TGF- β Family Signaling in the Control of Cell Proliferation and Survival. *Cold Spring Harbor perspectives in biology* **9**, a022145 (2017).
7. Kim, K.K., Sheppard, D. & Chapman, H.A. TGF- β 1 Signaling and Tissue Fibrosis. *Cold Spring Harbor perspectives in biology* **10**, a022293 (2017).
8. Goumans, M.J. & ten Dijke, P. TGF- β Signaling in Control of Cardiovascular Function. *Cold Spring Harbor perspectives in biology* **10**, a022210 (2017).
9. Colak, S. & ten Dijke, P. Targeting TGF- β Signaling in Cancer. *Trends in cancer* **3**, 56-71 (2017).
10. Akhurst, R.J. Targeting TGF- β Signaling for Therapeutic Gain. *Cold Spring Harbor perspectives in biology* **9**, a022301 (2017).
11. Derynck, R. *et al.* Human Transforming Growth Factor- β Complementary-DNA Sequence and Expression in Normal and Transformed-Cells. *Nature* **316**, 701-705 (1985).
12. Moses, H.L., Roberts, A.B. & Derynck, R. The discovery and early days of TGF- β : a historical perspective. *Cold Spring Harbor perspectives in biology* **8**, a021865 (2016).
13. Galat, A. Common structural traits for cystine knot domain of the TGF β superfamily of proteins and three-fingered ectodomain of their cellular receptors. *Cellular and molecular life sciences* **68**, 3437-3451 (2011).
14. Cheifetz, S. *et al.* The transforming growth factor- β system, a complex pattern of cross-reactive ligands and receptors. *Cell* **48**, 409-415 (1987).
15. Robertson, I.B. & Rifkin, D.B. Regulation of the Bioavailability of TGF- β and TGF- β -Related Proteins. *Cold Spring Harbor perspectives in biology* **8**, a021907 (2016).
16. Khan, Z. & Marshall, J.F. The role of integrins in TGF β activation in the tumour stroma. *Cell and tissue research* **365**, 657-673 (2016).
17. Wrana, J.L., Attisano, L., Wieser, R., Ventura, F. & Massagué, J. Mechanism of activation of the TGF- β receptor. *Nature* **370**, 341-347 (1994).
18. Yamashita, H., ten Dijke, P., Franzen, P., Miyazono, K. & Heldin, C.-H. Formation of hetero-oligomeric complexes of type I and type II receptors for transforming growth factor- β . *Journal of biological chemistry* **269**, 20172-20178 (1994).
19. López-Casillas, F., Wrana, J.L. & Massagué, J. Betaglycan presents ligand to the TGF β signaling receptor. *Cell* **73**, 1435-1444 (1993).
20. Villarreal, M.M. *et al.* Binding Properties of the Transforming Growth Factor- β Coreceptor Betaglycan: Proposed Mechanism for Potentiation of Receptor Complex Assembly and Signaling. *Biochemistry* **55**, 6880-6896 (2016).
21. ten Dijke, P. & Hill, C.S. New insights into TGF- β -SMAD signalling. *Trends in biochemical sciences* **29**, 265-273 (2004).
22. Shi, Y. & Massague, J. Mechanisms of TGF- β signaling from cell membrane to the nucleus. *Cell* **113**, 685-700 (2003).
23. Heldin, C.H., Miyazono, K. & ten Dijke, P. TGF- β signalling from cell membrane to nucleus through SMAD proteins. *Nature* **390**, 465-471 (1997).
24. Derynck, R. *et al.* Nomenclature: vertebrate mediators of TGF β family signals. *Cell* **87**, 173 (1996).
25. Nakao, A. *et al.* TGF- β receptor-mediated signalling through SMAD2, SMAD3 and SMAD4. *The EMBO journal* **16**, 5353-5362 (1997).
26. Nakao, A. *et al.* Identification of SMAD7, a TGF β -inducible antagonist of TGF- β signalling. *Nature* **389**, 631-635 (1997).
27. Hayashi, H. *et al.* The MAD-related protein SMAD7 associates with the TGF β receptor and functions as an antagonist of TGF β signaling. *Cell* **89**, 1165-1173 (1997).
28. Derynck, R. & Zhang, Y.E. SMAD-dependent and SMAD-independent pathways in TGF- β family

- signalling. *Nature* **425**, 577-584 (2003).
29. Zhang, Y.E. Non-SMAD Signaling Pathways of the TGF- β Family. *Cold Spring Harbor perspectives in biology* **9**, a022129 (2017).
 30. Mu, Y., Gudey, S.K. & Landstrom, M. Non-SMAD signaling pathways. *Cell and tissue research* **347**, 11-20 (2012).
 31. Lee, M.K. *et al.* TGF- β activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *The EMBO journal* **26**, 3957-3967 (2007).
 32. Arsur, M. *et al.* Transient activation of NF- κ B through a TAK1/IKK kinase pathway by TGF- β 1 inhibits AP-1/SMAD signaling and apoptosis: implications in liver tumor formation. *Oncogene* **22**, 412-425 (2003).
 33. Galliher, A.J. & Schiemann, W.P. Src phosphorylates Tyr284 in TGF- β type II receptor and regulates TGF- β stimulation of p38 MAPK during breast cancer cell proliferation and invasion. *Cancer Research* **67**, 3752-3758 (2007).
 34. Zavadil, J. *et al.* Genetic programs of epithelial cell plasticity directed by transforming growth factor- β . *Proceedings of the National Academy of Sciences of the United States of America* **98**, 6686-6691 (2001).
 35. Wang, C. *et al.* TAK1 is a ubiquitin-dependent kinase of MKK and IKK. *Nature* **412**, 346-351 (2001).
 36. Yi, J.Y., Shin, I. & Arteaga, C.L. Type I transforming growth factor β receptor binds to and activates phosphatidylinositol 3-kinase. *The Journal of biological chemistry* **280**, 10870-10876 (2005).
 37. Lamouille, S. & Derynck, R. Cell size and invasion in TGF- β -induced epithelial to mesenchymal transition is regulated by activation of the mTOR pathway. *Journal of cell biology* **178**, 437-451 (2007).
 38. Ozdamar, B. *et al.* Regulation of the polarity protein Par6 by TGF β receptors controls epithelial cell plasticity. *Science* **307**, 1603-1609 (2005).
 39. Wilkes, M.C., Murphy, S.J., Garamszegi, N. & Leof, E.B. Cell-type-specific activation of PAK2 by transforming growth factor β independent of SMAD2 and SMAD3. *Molecular and cellular biology* **23**, 8878-8889 (2003).
 40. Barrios-Rodiles, M. *et al.* High-throughput mapping of a dynamic signaling network in mammalian cells. *Science* **307**, 1621-1625 (2005).
 41. Robertson, I.B. & Rifkin, D.B. Regulation of the bioavailability of TGF- β and TGF- β -related proteins. *Cold Spring Harbor perspectives in biology* **8**, a021907 (2016).
 42. Shi, M. *et al.* Latent TGF- β structure and activation. *Nature* **474**, 343-349 (2011).
 43. Costanza, B., Umelo, I.A., Bellier, J., Castronovo, V. & Turtoi, A. Stromal Modulators of TGF- β in Cancer. *Journal of clinical medicine* **6**, 7 (2017).
 44. Robertson, I.B. *et al.* Latent TGF- β -binding proteins. *Matrix biology* **47**, 44-53 (2015).
 45. Isogai, Z. *et al.* Latent transforming growth factor- β -binding protein 1 interacts with fibrillin and is a microfibril-associated protein. *Journal of biological chemistry* **278**, 2750-2757 (2003).
 46. Mu, D. *et al.* The integrin α β 8 mediates epithelial homeostasis through MT1-MMP-dependent activation of TGF- β 1. *Journal of cell biology* **157**, 493-507 (2002).
 47. Yu, Q. & Stamenkovic, I. Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF- β and promotes tumor invasion and angiogenesis. *Genes & development* **14**, 163-176 (2000).
 48. Jobling, M.F. *et al.* Isoform-specific activation of latent transforming growth factor β (LTGF- β) by reactive oxygen species. *Radiation research* **166**, 839-848 (2006).
 49. Tran, D.Q. *et al.* GARP (LRRC32) is essential for the surface expression of latent TGF- β on platelets and activated FOXP3⁺ regulatory T cells. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13445-13450 (2009).
 50. Wang, R. *et al.* GARP regulates the bioavailability and activation of TGF β . *Molecular Biology of the Cell* **23**, 1129-1139 (2012).
 51. Annes, J.P., Chen, Y., Munger, J.S. & Rifkin, D.B. Integrin α V β 6-mediated activation of latent TGF- β requires the latent TGF- β binding protein-1. *Journal of cell biology* **165**, 723-734 (2004).
 52. Dong, X. *et al.* Force interacts with macromolecular structure in activation of TGF- β . *Nature* **542**, 55-59 (2017).
 53. Munger, J.S. & Sheppard, D. Cross talk among TGF- β signaling pathways, integrins, and the extracellular matrix. *Cold Spring Harbor perspectives in biology* **3**, a005017 (2011).
 54. Dallas, S.L. *et al.* Fibronectin regulates latent transforming growth factor- β (TGF β) by controlling matrix assembly of latent TGF β -binding protein-1. *The Journal of biological chemistry* **280**, 18871-18880 (2005).
 55. Merline, R. *et al.* Signaling by the matrix proteoglycan decorin controls inflammation and cancer through PDCD4 and MicroRNA-21. *Science signaling* **4**, ra75 (2011).
 56. Goetschy, J.F., Letourneur, O., Cerletti, N. & Horisberger, M.A. The unglycosylated extracellular domain

- of type-II receptor for transforming growth factor- β . A novel assay for characterizing ligand affinity and specificity. *European journal of biochemistry* **241**, 355-362 (1996).
57. Kim, Y.W., Park, J., Lee, H.J., Lee, S.Y. & Kim, S.J. TGF- β sensitivity is determined by N-linked glycosylation of the type II TGF- β receptor. *The Biochemical journal* **445**, 403-411 (2012).
 58. Ebisawa, T. *et al.* Smurf1 interacts with transforming growth factor- β type I receptor through SMAD7 and induces receptor degradation. *The Journal of biological chemistry* **276**, 12477-12480 (2001).
 59. Kavsak, P. *et al.* SMAD7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF β receptor for degradation. *Molecular cell* **6**, 1365-1375 (2000).
 60. Liu, S., de Boeck, M., van Dam, H. & ten Dijke, P. Regulation of the TGF- β pathway by deubiquitinases in cancer. *The international journal of biochemistry & cell biology* **76**, 135-145 (2016).
 61. Shi, W. *et al.* GADD34-PP1c recruited by SMAD7 dephosphorylates TGF β type I receptor. *Journal cell biology* **164**, 291-300 (2004).
 62. Batut, J. *et al.* Two highly related regulatory subunits of PP2A exert opposite effects on TGF- β /Activin/Nodal signalling. *Development* **135**, 2927-2937 (2008).
 63. Budi, E.H., Duan, D. & Derynck, R. Transforming Growth Factor- β Receptors and SMADs: Regulatory Complexity and Functional Versatility. *Trends in cell biology* (2017).
 64. Stenvers, K.L. *et al.* Heart and liver defects and reduced transforming growth factor β 2 sensitivity in transforming growth factor β type III receptor-deficient embryos. *Molecular and cell biology* **23**, 4371-4385 (2003).
 65. Lebrin, F. *et al.* Endoglin promotes endothelial cell proliferation and TGF- β /ALK1 signal transduction. *The EMBO journal* **23**, 4018-4028 (2004).
 66. Scherner, O., Meurer, S.K., Tihaa, L., Gressner, A.M. & Weiskirchen, R. Endoglin differentially modulates antagonistic transforming growth factor- β 1 and BMP-7 signaling. *The Journal of biological chemistry* **282**, 13934-13943 (2007).
 67. Chen, H.B., Shen, J., Ip, Y.T. & Xu, L. Identification of phosphatases for SMAD in the BMP/DPP pathway. *Genes & development* **20**, 648-653 (2006).
 68. Sapkota, G. *et al.* Dephosphorylation of the linker regions of SMAD1 and SMAD2/3 by small C-terminal domain phosphatases has distinct outcomes for bone morphogenetic protein and transforming growth factor- β pathways. *The Journal of biological chemistry* **281**, 40412-40419 (2006).
 69. Lin, X. *et al.* PPM1A functions as a SMAD phosphatase to terminate TGF β signaling. *Cell* **125**, 915-928 (2006).
 70. Lo, R.S. & Massague, J. Ubiquitin-dependent degradation of TGF- β -activated SMAD2. *Nature cell biology* **1**, 472-478 (1999).
 71. Gao, S. *et al.* Ubiquitin ligase Nedd4L targets activated SMAD2/3 to limit TGF- β signaling. *Molecular cell* **36**, 457-468 (2009).
 72. Zhou, F. *et al.* USP4 inhibits SMAD4 monoubiquitination and promotes activin and BMP signaling. *The EMBO journal* (2017).
 73. Gulei, D. *et al.* The "good-cop bad-cop" TGF- β role in breast cancer modulated by non-coding RNAs. *Biochimica et biophysica acta* **1861**, 1661-1675 (2017).
 74. Chen, Y. *et al.* miR-200b inhibits TGF- β 1-induced epithelial-mesenchymal transition and promotes growth of intestinal epithelial cells. *Cell death & disease* **4**, e541 (2013).
 75. Itoh, S., Itoh, F., Goumans, M.J. & ten Dijke, P. Signaling of transforming growth factor- β family members through SMAD proteins. *European journal of biochemistry* **267**, 6954-6967 (2000).
 76. Piek, E., Heldin, C.H. & ten Dijke, P. Specificity, diversity, and regulation in TGF- β superfamily signaling. *FASEB Journal* **13**, 2105-2124 (1999).
 77. Mullen, A.C. *et al.* Master transcription factors determine cell-type-specific responses to TGF- β signaling. *Cell* **147**, 565-576 (2011).
 78. Davis, B.N., Hilyard, A.C., Lagna, G. & Hata, A. SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* **454**, 56-61 (2008).
 79. Mondal, T. *et al.* MEG3 long noncoding RNA regulates the TGF- β pathway genes through formation of RNA-DNA triplex structures. *Nature communications* **6**, 7743 (2015).
 80. Schuster, N. & Kriegelstein, K. Mechanisms of TGF- β -mediated apoptosis. *Cell and tissue research* **307**, 1-14 (2002).
 81. Macias, M.J., Martin-Malpartida, P. & Massague, J. Structural determinants of SMAD function in TGF- β signaling. *Trends in biochemical sciences* **40**, 296-308 (2015).
 82. Fabregat, I., Fernando, J., Mainez, J. & Sancho, P. TGF- β signaling in cancer treatment. *Current pharmaceutical design* **20**, 2934-2947 (2014).
 83. Chandrasinghe, P. *et al.* Role of SMAD proteins in colitis-associated cancer: from known to the unknown.

- Oncogene* **37**, 1-7 (2017).
84. ten Dijke, P., Goumans, M.J., Itoh, F. & Itoh, S. Regulation of cell proliferation by SMAD proteins. *Journal of cellular physiology* **191**, 1-16 (2002).
 85. Liu, F. Delineating the TGF- β /SMAD-Induced Cytostatic Response. *SMAD signal transduction* 75-91 (2006).
 86. Massague, J., Blain, S.W. & Lo, R.S. TGF β signaling in growth control, cancer, and heritable disorders. *Cell* **103**, 295-309 (2000).
 87. Tarasewicz, E. & Jeruss, J.S. Phospho-specific SMAD3 signaling: impact on breast oncogenesis. *Cell cycle* **11**, 2443-2451 (2012).
 88. Chen, C.-R., Kang, Y., Siegel, P.M. & Massagué, J. E2F4/5 and p107 as SMAD cofactors linking the TGF β receptor to c-myc repression. *Cell* **110**, 19-32 (2002).
 89. Siegel, P.M. & Massague, J. Cytostatic and apoptotic actions of TGF- β in homeostasis and cancer. *Nature reviews Cancer* **3**, 807-821 (2003).
 90. Ozaki, I., Hamajima, H., Matsushashi, S. & Mizuta, T. Regulation of TGF- β 1-Induced Pro-Apoptotic Signaling by Growth Factor Receptors and Extracellular Matrix Receptor Integrins in the Liver. *Frontiers in physiology* **2**, 78 (2011).
 91. Moustakas, A. & Heldin, C.H. Non-SMAD TGF- β signals. *Journal of cell science* **118**, 3573-3584 (2005).
 92. Perlman, R., Schiemann, W.P., Brooks, M.W., Lodish, H.F. & Weinberg, R.A. TGF- β -induced apoptosis is mediated by the adapter protein Daxx that facilitates JNK activation. *Nature cell biology* **3**, 708-714 (2001).
 93. Jang, C.W. *et al.* TGF- β induces apoptosis through SMAD-mediated expression of DAP-kinase. *Nature cell biology* **4**, 51-58 (2002).
 94. Kokkinakis, D.M. *et al.* Modulation of gene expression in human central nervous system tumors under methionine deprivation-induced stress. *Cancer research* **64**, 7513-7525 (2004).
 95. Wang, J. *et al.* Transforming growth factor β induces apoptosis through repressing the phosphoinositide 3-kinase/AKT/survivin pathway in colon cancer cells. *Cancer research* **68**, 3152-3160 (2008).
 96. Welm, A.L. TGF β primes breast tumor cells for metastasis. *Cell* **133**, 27-28 (2008).
 97. Massague, J. TGF β in Cancer. *Cell* **134**, 215-230 (2008).
 98. Derynck, R., Akhurst, R.J. & Balmain, A. TGF- β signaling in tumor suppression and cancer progression. *Nature genetics* **29**, 117-129 (2001).
 99. Alvi, A.J., Rader, J.S., Brogini, M., Latif, F. & Maher, E.R. Microsatellite instability and mutational analysis of transforming growth factor β receptor type II gene (TGFBR2) in sporadic ovarian cancer. *Molecular pathology* **54**, 240-243 (2001).
 100. Nagata, H. *et al.* Inhibition of c-Jun NH2-terminal kinase switches SMAD3 signaling from oncogenesis to tumor-suppression in rat hepatocellular carcinoma. *Hepatology* **49**, 1944-1953 (2009).
 101. Meulmeester, E. & ten Dijke, P. The dynamic roles of TGF- β in cancer. *The Journal of pathology* **223**, 205-218 (2011).
 102. Seoane, J. & Gomis, R.R. TGF- β Family Signaling in Tumor Suppression and Cancer Progression. *Cold Spring Harbor perspectives in biology* (2017).
 103. Drabsch, Y. & ten Dijke, P. TGF- β signalling and its role in cancer progression and metastasis. *Cancer metastasis reviews* **31**, 553-568 (2012).
 104. Huang, J.J. & Blobel, G.C. Dichotomous roles of TGF- β in human cancer. *Biochemical society transactions* **44**, 1441-1454 (2016).
 105. Thiery, J.P., Acloque, H., Huang, R.Y. & Nieto, M.A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871-890 (2009).
 106. Nieto, M.A., Huang, R.Y., Jackson, R.A. & Thiery, J.P. EMT: 2016. *Cell* **166**, 21-45 (2016).
 107. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nature reviews Molecular cell biology* **15**, 178-196 (2014).
 108. Comijn, J. *et al.* The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Molecular cell* **7**, 1267-1278 (2001).
 109. Vincent, T. *et al.* A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF- β mediated epithelial-mesenchymal transition. *Nature cell biology* **11**, 943-950 (2009).
 110. Kong, W. *et al.* MicroRNA-155 is regulated by the transforming growth factor β /SMAD pathway and contributes to epithelial cell plasticity by targeting RhoA. *Molecular and cellular biology* **28**, 6773-6784 (2008).
 111. Johansson, J. *et al.* MiR-155-mediated loss of C/EBP β shifts the TGF- β response from growth inhibition to epithelial-mesenchymal transition, invasion and metastasis in breast cancer. *Oncogene* **32**, 5614-5624 (2013).
 112. Yuan, J.H. *et al.* A long noncoding RNA activated by TGF- β promotes the invasion-metastasis cascade

- in hepatocellular carcinoma. *Cancer cell* **25**, 666-681 (2014).
113. Moustakas, A. & Heldin, C.H. Mechanisms of TGF β -Induced Epithelial-Mesenchymal Transition. *Journal of clinical medicine* **5**,63 (2016).
 114. Janda, E. *et al.* Ras and TGF β cooperatively regulate epithelial cell plasticity and metastasis. *Journal of cell biology* **156**, 299-314 (2002).
 115. Zhang, L., Zhou, F. & ten Dijke, P. Signaling interplay between transforming growth factor- β receptor and PI3K/AKT pathways in cancer. *Trends in biochemical sciences* **38**, 612-620 (2013).
 116. Bakin, A.V., Rinehart, C., Tomlinson, A.K. & Arteaga, C.L. p38 mitogen-activated protein kinase is required for TGF β -mediated fibroblastic transdifferentiation and cell migration. *Journal of cell Science* **115**, 3193-3206 (2002).
 117. Yamashita, M. *et al.* TRAF6 mediates SMAD-independent activation of JNK and p38 by TGF- β . *Molecular cell* **31**, 918-924 (2008).
 118. Zhang, L. *et al.* TRAF4 promotes TGF- β receptor signaling and drives breast cancer metastasis. *Molecular cell* **51**, 559-572 (2013).
 119. Kretzschmar, M., Doody, J., Timokhina, I. & Massagué, J. A mechanism of repression of TGF β /SMAD signaling by oncogenic Ras. *Genes & development* **13**, 804-816 (1999).
 120. Zhou, B.P. *et al.* Dual regulation of Snail by GSK-3 β -mediated phosphorylation in control of epithelial-mesenchymal transition. *Nature cell biology* **6**, 931-940 (2004).
 121. Julien, S. *et al.* Activation of NF- κ B by Akt upregulates Snail expression and induces epithelium mesenchyme transition. *Oncogene* **26**, 7445-7456 (2007).
 122. Xue, G. *et al.* Akt/PKB-mediated phosphorylation of Twist1 promotes tumor metastasis via mediating cross-talk between PI3K/Akt and TGF- β signaling axes. *Cancer discovery* **2**, 248-259 (2012).
 123. Bhowmick, N.A. *et al.* Transforming growth factor- β 1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism. *Molecular biology of the cell* **12**, 27-36 (2001).
 124. Menezes, M.E. *et al.* MDA-9/Syntenin (SDCBP) modulates small GTPases RhoA and Cdc42 via transforming growth factor β 1 to enhance epithelial-mesenchymal transition in breast cancer. *Oncotarget* **7**, 80175-80189 (2016).
 125. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
 126. Wiercinska, E. *et al.* The TGF- β /SMAD pathway induces breast cancer cell invasion through the up-regulation of matrix metalloproteinase 2 and 9 in a spheroid invasion model system. *Breast cancer research and treatment* **128**, 657-666 (2011).
 127. Stuelten, C.H. *et al.* Breast cancer cells induce stromal fibroblasts to express MMP-9 via secretion of TNF- α and TGF- β . *Journal of cell science* **118**, 2143-2153 (2005).
 128. Wang, B. *et al.* TGF β -mediated upregulation of hepatic miR-181b promotes hepatocarcinogenesis by targeting TIMP3. *Oncogene* **29**, 1787-1797 (2010).
 129. Valastyan, S. & Weinberg, R.A. Tumor metastasis: molecular insights and evolving paradigms. *Cell* **147**, 275-292 (2011).
 130. Mundy, G.R. Metastasis to bone: causes, consequences and therapeutic opportunities. *Nature reviews Cancer* **2**, 584-593 (2002).
 131. Suva, L.J., Washam, C., Nicholas, R.W. & Griffin, R.J. Bone metastasis: mechanisms and therapeutic opportunities. *Nature reviews Endocrinology* **7**, 208-218 (2011).
 132. Yin, J.J. *et al.* TGF- β signaling blockade inhibits PTHrP secretion by breast cancer cells and bone metastases development. *The Journal of clinical investigation* **103**, 197-206 (1999).
 133. Kakonen, S.M. *et al.* Transforming growth factor- β stimulates parathyroid hormone-related protein and osteolytic metastases via SMAD and mitogen-activated protein kinase signaling pathways. *The Journal of biological chemistry* **277**, 24571-24578 (2002).
 134. Lacey, D.L. *et al.* Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation. *Cell* **93**, 165-176 (1998).
 135. Kang, Y. *et al.* A multigenic program mediating breast cancer metastasis to bone. *Cancer cell* **3**, 537-549 (2003).
 136. Kang, Y. *et al.* Breast cancer bone metastasis mediated by the SMAD tumor suppressor pathway. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 13909-13914 (2005).
 137. Zheng, H., Li, W. & Kang, Y. Tumor-Stroma Interactions in Bone Metastasis: Molecular Mechanisms and Therapeutic Implications. *Cold Spring Harbor symposia on quantitative biology* **81**, 151-161 (2016).
 138. Padua, D. *et al.* TGF β primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**, 66-77 (2008).
 139. Kang, N., Gores, G.J. & Shah, V.H. Hepatic stellate cells: partners in crime for liver metastases?

- Hepatology* **54**, 707-713 (2011).
140. Tu, K. *et al.* Vasodilator-stimulated phosphoprotein promotes activation of hepatic stellate cells by regulating Rab11-dependent plasma membrane targeting of transforming growth factor β receptors. *Hepatology* **61**, 361-374 (2015).
 141. Liu, C. *et al.* IQGAP1 suppresses T β RII-mediated myofibroblastic activation and metastatic growth in liver. *The Journal of clinical investigation* **123**, 1138-1156 (2013).
 142. Carmeliet, P. & Jain, R.K. Angiogenesis in cancer and other diseases. *Nature* **407**, 249-257 (2000).
 143. Ito, N. *et al.* Positive correlation of plasma transforming growth factor- β 1 levels with tumor vascularity in hepatocellular carcinoma. *Cancer letters* **89**, 45-48 (1995).
 144. Ivanovic, V., Melman, A., Davis-Joseph, B., Valcic, M. & Geliebter, J. Elevated plasma levels of TGF- β 1 in patients with invasive prostate cancer. *Nature medicine* **1**, 282-284 (1995).
 145. Wikstrom, P., Stattin, P., Franck-Lissbrant, I., Damber, J.E. & Bergh, A. Transforming growth factor β 1 is associated with angiogenesis, metastasis, and poor clinical outcome in prostate cancer. *The Prostate* **37**, 19-29 (1998).
 146. Hasegawa, Y. *et al.* Transforming growth factor- β 1 level correlates with angiogenesis, tumor progression, and prognosis in patients with nonsmall cell lung carcinoma. *Cancer* **91**, 964-971 (2001).
 147. de Vinuesa, A.G., Bocci, M., Pietras, K. & ten Dijke, P. Targeting tumour vasculature by inhibiting activin receptor-like kinase (ALK) 1 function. *Biochemical Society transactions* **44**, 1142-1149 (2016).
 148. Paauwe, M., ten Dijke, P. & Hawinkels, L.J. Endoglin for tumor imaging and targeted cancer therapy. *Expert opinion on therapeutic targets* **17**, 421-435 (2013).
 149. Pertovaara, L. *et al.* Vascular endothelial growth factor is induced in response to transforming growth factor- β in fibroblastic and epithelial cells. *The Journal of biological chemistry* **269**, 6271-6274 (1994).
 150. Sanchez-Elsner, T. *et al.* Synergistic cooperation between hypoxia and transforming growth factor- β pathways on human vascular endothelial growth factor gene expression. *The Journal of biological chemistry* **276**, 38527-38535 (2001).
 151. Petersen, M. *et al.* SMAD2 and SMAD3 have opposing roles in breast cancer bone metastasis by differentially affecting tumor angiogenesis. *Oncogene* **29**, 1351-1361 (2010).
 152. Thomas, D.A. & Massague, J. TGF- β directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. *Cancer cell* **8**, 369-380 (2005).
 153. Ahmadzadeh, M. & Rosenberg, S.A. TGF- β 1 attenuates the acquisition and expression of effector function by tumor antigen-specific human memory CD8 T cells. *Journal of immunology* **174**, 5215-5223 (2005).
 154. Fu, S. *et al.* TGF- β induces Foxp3 + T-regulatory cells from CD4 + CD25 - precursors. *American journal of transplantation* **4**, 1614-1627 (2004).
 155. Liu, Y. *et al.* A critical function for TGF- β signaling in the development of natural CD4+CD25+Foxp3+ regulatory T cells. *Nature immunology* **9**, 632-640 (2008).
 156. Somasundaram, R. *et al.* Inhibition of cytolytic T lymphocyte proliferation by autologous CD4+/CD25+ regulatory T cells in a colorectal carcinoma patient is mediated by transforming growth factor- β . *Cancer research* **62**, 5267-5272 (2002).
 157. Wilson, E.B. *et al.* Human tumour immune evasion via TGF- β blocks NK cell activation but not survival allowing therapeutic restoration of anti-tumour activity. *PLoS One* **6**, e22842 (2011).
 158. Crane, C.A. *et al.* TGF- β downregulates the activating receptor NKG2D on NK cells and CD8+ T cells in glioma patients. *Neuro-oncology* **12**, 7-13 (2010).
 159. Donatelli, S.S. *et al.* TGF- β -inducible microRNA-183 silences tumor-associated natural killer cells. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4203-4208 (2014).
 160. Sica, A., Schioppa, T., Mantovani, A. & Allavena, P. Tumour-associated macrophages are a distinct M2 polarised population promoting tumour progression: potential targets of anti-cancer therapy. *European journal of cancer* **42**, 717-727 (2006).
 161. Gong, D. *et al.* TGF β signaling plays a critical role in promoting alternative macrophage activation. *BMC immunology* **13**, 31 (2012).
 162. Crooke, S.T. Molecular mechanisms of action of antisense drugs. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* **1489**, 31-43 (1999).
 163. Bogdahn, U. *et al.* Results of a phase I/II active-controlled study with AP 12009 for patients with recurrent or refractory anaplastic astrocytoma. *Journal of clinical oncology* **26**, 2076-2076 (2008).
 164. Hau, P. *et al.* Inhibition of TGF- β 2 with ap 12009 in recurrent malignant gliomas: from preclinical to phase I/II studies. *Oligonucleotides* **17**, 201-212 (2007).
 165. Nagaraj, N.S. & Datta, P.K. Targeting the transforming growth factor- β signaling pathway in human cancer. *Expert opinion on investigational drugs* **19**, 77-91 (2010).
 166. Bogdahn, U. *et al.* Targeted therapy for high-grade glioma with the TGF- β 2 inhibitor trabedersen: results

- of a randomized and controlled phase IIb study. *Neuro-oncology* **13**, 132-142 (2010).
167. Schlingensiepen, K.-H. *et al.* The TGF- β 1 antisense oligonucleotide AP 11014 for the treatment of non-small cell lung, colorectal and prostate cancer: Preclinical studies. *Journal of clinical oncology* **22**, 3132-3132 (2004).
168. Mohammad, K.S. *et al.* TGF- β -RI kinase inhibitor SD-208 reduces the development and progression of melanoma bone metastases. *Cancer research* **71**, 175-184 (2011).
169. Gaspar, N.J. *et al.* Inhibition of transforming growth factor β signaling reduces pancreatic adenocarcinoma growth and invasiveness. *Molecular pharmacology* **72**, 152-161 (2007).
170. Ge, R. *et al.* Inhibition of growth and metastasis of mouse mammary carcinoma by selective inhibitor of transforming growth factor- β type I receptor kinase in vivo. *Clinical cancer research* **12**, 4315-4330 (2006).
171. Bueno, L. *et al.* Semi-mechanistic modelling of the tumour growth inhibitory effects of LY2157299, a new type I receptor TGF- β kinase antagonist, in mice. *European journal of cancer* **44**, 142-150 (2008).
172. Herbertz, S. *et al.* Clinical development of galunisertib (LY2157299 monohydrate), a small molecule inhibitor of transforming growth factor- β signaling pathway. *Drug design, development and therapy* **9**, 4479 (2015).
173. Rodon, J. *et al.* First-in-human dose study of the novel transforming growth factor- β receptor I kinase inhibitor LY2157299 monohydrate in patients with advanced cancer and glioma. *Clinical cancer research* **21**, 553-560 (2014).
174. Zhang, B. *et al.* Antimetastatic role of SMAD4 signaling in colorectal cancer. *Gastroenterology* **138**, 969-980. e963 (2010).
175. Melisi, D. *et al.* LY2109761, a novel transforming growth factor β receptor type I and type II dual inhibitor, as a therapeutic approach to suppressing pancreatic cancer metastasis. *Molecular cancer therapeutics* **7**, 829-840 (2008).
176. Zhang, M. *et al.* Blockade of TGF- β signaling by the TGF β R-I kinase inhibitor LY2109761 enhances radiation response and prolongs survival in glioblastoma. *Cancer research* **71**, 7155-7167 (2011).
177. Alsina-Sanchis, E. *et al.* The TGF β pathway stimulates ovarian cancer cell proliferation by increasing IGF1R levels. *International journal of cancer* **139**, 1894-1903 (2016).
178. Gao, Y. *et al.* LY2109761 enhances cisplatin antitumor activity in ovarian cancer cells. *International journal of clinical and experimental pathology* **8**, 4923 (2015).
179. Kim, B.G. *et al.* TGF- β type I receptor inhibitor (TEW-7197) diminishes myeloma progression by multiple immunomodulatory mechanisms in combination with ixazomib. *Cancer research* **77**, 2647 (2017).
180. Colak, S. & ten Dijke, P. Targeting TGF- β Signaling in Cancer. *Trends in Cancer* **3**, 56-71 (2017).
181. Nam, J.-S. *et al.* An anti-transforming growth factor β antibody suppresses metastasis via cooperative effects on multiple cell compartments. *Cancer research* **68**, 3835-3843 (2008).
182. Sturge, J., Caley, M.P. & Waxman, J. Bone metastasis in prostate cancer: emerging therapeutic strategies. *Nature reviews Clinical oncology* **8**, 357-368 (2011).
183. Kobie, J.J. *et al.* Transforming growth factor β inhibits the antigen-presenting functions and antitumor activity of dendritic cell vaccines. *Cancer Research* **63**, 1860-1864 (2003).
184. Nam, J.-S. *et al.* Bone sialoprotein mediates the tumor cell-targeted prometastatic activity of transforming growth factor β in a mouse model of breast cancer. *Cancer research* **66**, 6327-6335 (2006).
185. Morris, J.C. *et al.* Phase I study of GC1008 (fresolimumab): a human anti-transforming growth factor- β (TGF β) monoclonal antibody in patients with advanced malignant melanoma or renal cell carcinoma. *PLoS one* **9**, e90353 (2014).
186. Muraoka, R.S. *et al.* Blockade of TGF- β inhibits mammary tumor cell viability, migration, and metastases. *The Journal of clinical investigation* **109**, 1551 (2002).
187. Suzuki, E. *et al.* Soluble type II transforming growth factor- β receptor inhibits established murine malignant mesothelioma tumor growth by augmenting host antitumor immunity. *Clinical cancer research* **10**, 5907-5918 (2004).
188. Bandyopadhyay, A. *et al.* Antitumor activity of a recombinant soluble betaglycan in human breast cancer xenograft. *Cancer research* **62**, 4690-4695 (2002).
189. Finger, E.C. *et al.* T β RIII suppresses non-small cell lung cancer invasiveness and tumorigenicity. *Carcinogenesis* **29**, 528-535 (2008).
190. Naumann, U. *et al.* Glioma gene therapy with soluble transforming growth factor- β receptors II and III. *International journal of oncology* **33**, 759-765 (2008).
191. Huntley, S.P. *et al.* Attenuated type II TGF- β receptor signalling in human malignant oral keratinocytes induces a less differentiated and more aggressive phenotype that is associated with metastatic dissemination. *International journal of cancer* **110**, 170-176 (2004).

192. Tang, B. *et al.* TGF- β switches from tumor suppressor to prometastatic factor in a model of breast cancer progression. *Journal of clinical investigation* **112**, 1116 (2003).

2) Long non-coding RNAs in TGF- β signaling and EMT

Chuannan Fan[#], Qian Wang[#], and Peter ten Dijke*

Oncode Institute, Leiden University Medical Center, Leiden, the Netherlands.

[#] These authors contributed equally.

*Correspondence: Peter ten Dijke

Department of Cell and Chemical Biology and Oncode Institute, Leiden University Medical Center, Einthovenweg 20, 2300 RC, Leiden, The Netherlands

Email: p.ten_dijke@lumc.nl; Telephone: +31 71526 9271; Fax: +31 71 526 8270

Abstract

Transforming growth factor- β (TGF- β) signaling can have a dual role during cancer progression and suppress tumorigenesis at initial stages of cancer but promote cancer progression at advanced stages. The latter is achieved, in part, by acting directly on cancer cells by inducing a transition from epithelial to a highly invasive mesenchymal state (EMT). Ligand-induced activation of transmembrane TGF- β receptor triggers EMT through activation of intracellular SMAD transcription factors. TGF- β signaling is regulated by modulators at multiple levels during EMT. Although the importance of protein coding genes that are modulated in response to TGF- β /SMAD signaling have been well studied, an important role of long non-coding RNAs (lncRNAs) in TGF- β /SMAD signaling action is emerging. This mini-review focusses on the mechanisms by which lncRNAs interplay with TGF- β signaling.

Molecular Basis of lncRNAs

Although more than 70% of human genome can be actively transcribed, only around 2% of it is transcribed into protein coding messenger RNAs (mRNAs)^{1, 2}. However, a large amount of lncRNAs, which had been recognized as “transcription noise” for a long time, is extensively transcribed within the human genome^{3, 4}. A recent study that collected the sequencing results from various publicly available databases revealed 95,243 human lncRNA genes and 323,950 human lncRNA transcripts⁵. lncRNAs are arbitrarily defined by that their transcript length is longer than 200 nucleotides (nt). Similar to mRNAs, most lncRNAs are transcribed by RNA polymerase II (Pol II), and a large proportion of lncRNAs undergo alternative splicing and polyadenylation^{6, 7}. Unlike mRNAs, lncRNA primary sequences are less conserved among species and lncRNA expression generally exhibits high tissue specificity^{2, 8-10}.

Mechanisms of lncRNAs

lncRNAs can be divided into nuclear and cytoplasmic lncRNAs depending on their subcellular localization (Fig. 1). By interacting with chromatin modifiers or transcription (co)factors, nuclear lncRNAs can alter the epigenetic landscape or the transcription process, and thereby change target gene expression¹¹⁻¹⁵. Nuclear lncRNAs can influence RNA splicing by interacting with the serine and arginine-rich (SR) protein^{16, 17}. Moreover, a subgroup of lncRNAs called enhancer RNAs, which are transcribed from active enhancers, can modulate chromatin looping *in cis* or *in trans*, leading to the activation of target gene transcription¹⁸⁻²¹. Cytoplasmic lncRNAs can regulate mRNA stability or translation through directly binding to mRNAs or RNA binding proteins (RBPs). lncRNAs localized in the cytoplasm can also act as sponges

for microRNAs (miRNAs)^{22, 23}. Recent studies have shown that functional small peptides can be encoded by cytoplasmic lncRNAs that associate with ribosomes^{24, 25}. Additionally, both cytoplasmic and nuclear lncRNAs can regulate protein post-translational modifications (PTMs) or molecular complex formation by functioning as scaffolds or decoys²⁶⁻²⁹.

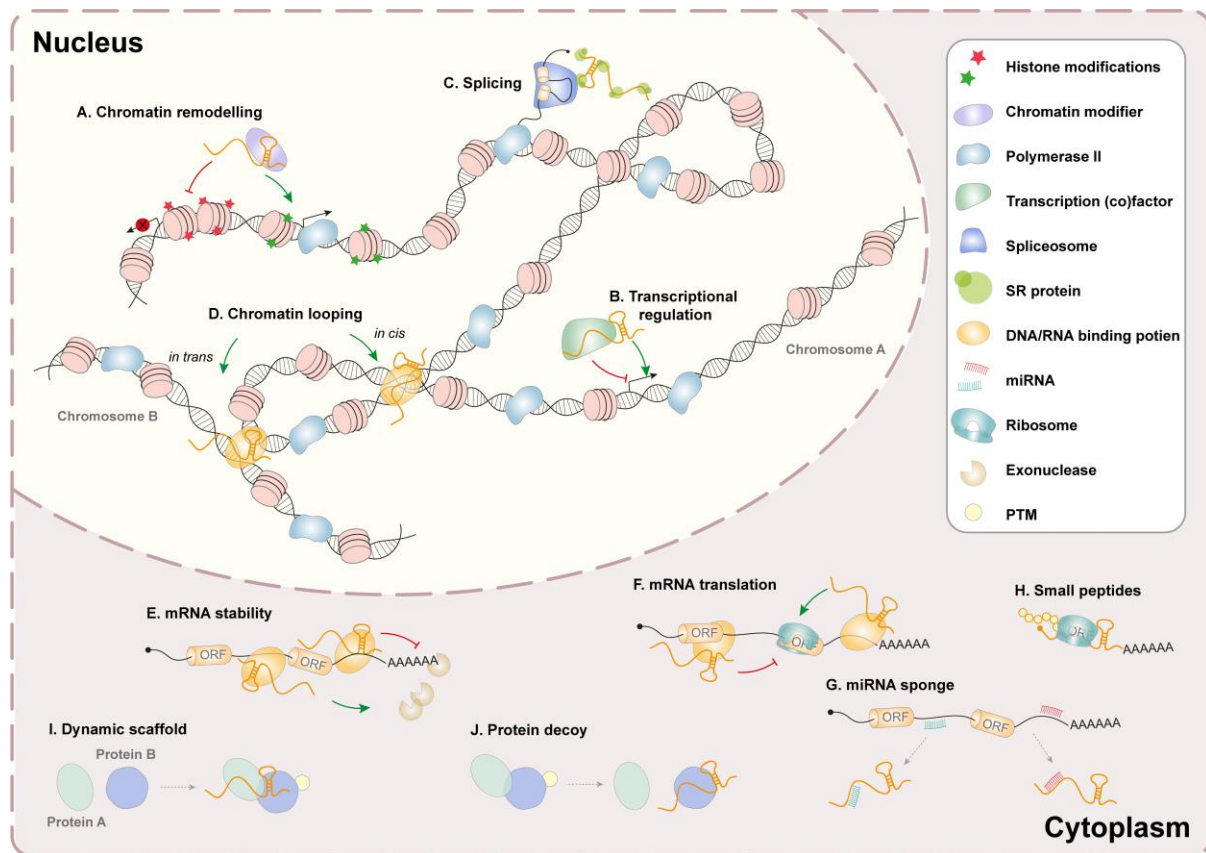


Fig.1 Mechanisms of lncRNAs function. Nuclear lncRNAs can interact with chromatin modifiers (A) or transcription (co)factors (B) to regulate chromatin landscape or gene transcription. They can also mediate alternative splicing (C) and chromatin looping (D). Cytoplasmic lncRNAs can affect mRNA stability (E) or translation (F), sponge miRNAs (G), encode small peptides (H), and modulate protein interactions and post-translational modification (I, J). ORF: open reading frame; SR: serine and arginine-rich; PTM: post-translational modification.

LncRNAs Function as Effectors of TGF- β Signaling

TGF- β -induced gene products frequently function as effectors of TGF- β -induced responses, for example EMT^{30, 31}. Consistent with this scenario, TGF- β -induced lncRNAs can drive TGF- β -induced EMT in cancer. *LncRNA-HOXA transcript induced by TGF- β (LncRNA-HIT)* promotes TGF- β -induced EMT and migration by specifically mitigating E-cadherin expression in mouse mammary NMuMG cells³². TGF- β promotes the expression of *lncRNA-activated by TGF- β (lncRNA-ATB)*, which stabilizes *interleukin-11 (IL-11)* mRNA, resulting in the promotion of hepatocellular carcinoma (HCC) cell colonization in secondary tissues³³. In addition, *lncRNA-ATB* drives EMT by serving as a sponge for *miR-200*, leading to the upregulation of EMT transcription factor ZEB1/2³³. Moreover, expression of other TGF- β downstream EMT transcription factors including SNAIL^{34, 35}, SLUG³⁴⁻³⁶ and TWIST^{37, 38} can be activated by TGF- β -induced lncRNAs.

TGF- β can induce lncRNA expression to influence the transcriptional output by altering epigenetic modifications. TGF- β -induced *Metastasis Associated Lung Adenocarcinoma Transcript 1 (MALAT1)* interacts with H3K27 methyltransferase suppressor of zeste 12 (suz12),

a component of the polycomb repressive complex 2 (PRC2), to promote H3K27me3 abundance at the promoter of *CDH1* (the gene that encodes E-cadherin) and to potentiate EMT in bladder cancer cells³⁹. *TGFB2-antisense RNA1* (*TGFB2-AS1*) associates with PRC2 adaptor protein EED to facilitate H3K27me3 modification at the promoter of TGF- β target genes⁴⁰.

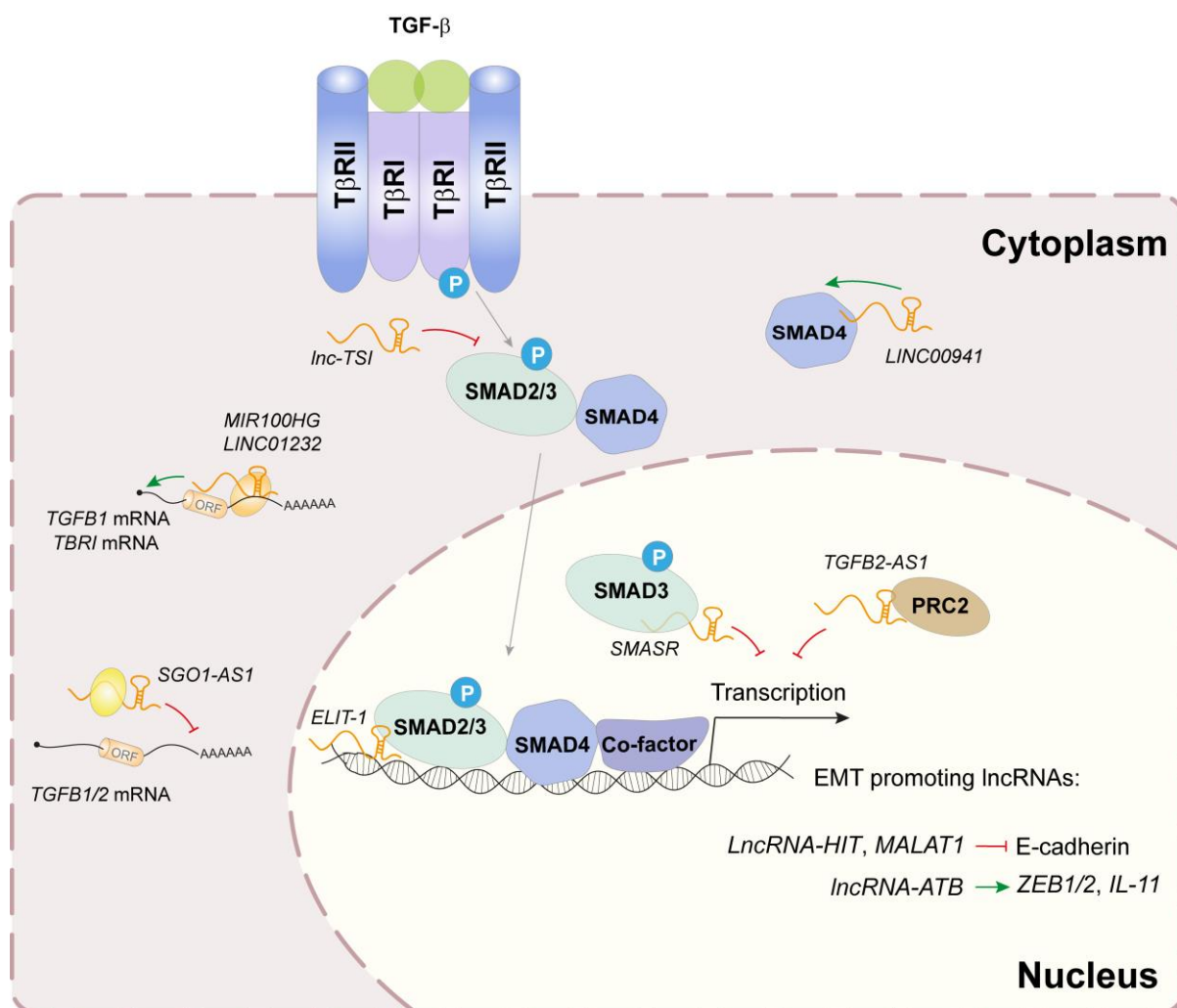


Fig.2 Interplay between lncRNAs and TGF- β signaling. TGF- β signaling induces lncRNAs to regulate EMT in cancer. lncRNAs can also modulate TGF- β signaling transduction at different levels, from ligand production to transcriptional output. (For description see text)

LncRNAs Function as Modulators of TGF- β Signaling

LncRNAs can act as modulators to fine-tune TGF- β signaling transduction in a negative or positive feedback manner^{41, 42}. TGF- β induced *mir-100-let-7a-2-mir-125b-1 cluster host gene* (*MIR100HG*) enhances *TGFB1* mRNA stability by promoting the binding of RBP HuR to *TGFB1* mRNA in multiple cancer cells⁴³. *MIR100HG* enhances TGF- β 1 autocrine to potentiate TGF- β signaling⁴³. *SGO1-AS1* facilitates *TGFB1/2* mRNA decay by competing their binding to PTBP1, an RBP that stabilizes *TGFB1/2* mRNA⁴⁴. TGF- β 1/2 production is therefore decreased by *SGO1-AS1*, leading to the attenuation of EMT and cancer metastasis⁴⁴.

Expression of TGF- β signaling receptors is regulated by lncRNAs. *SMAD3-associated long non-coding RNA* (*SMASR*) expression is suppressed by TGF- β /SMAD signaling in lung adenocarcinoma cells⁴⁵. *SMASR* interacts with SMAD3 to attenuate *TBRI* mRNA transcription, thus leading to inactivation of TGF- β /SMAD signaling⁴⁵. *LINC01232* recruits the RBP insulin

like growth factor 2 mRNA binding protein 2 (IGF2BP2) to protect *TBRI* mRNA from degradation⁴⁶. As a consequence, TGF- β signaling and cell stemness are potentiated by *LINC01232* in lung adenocarcinoma cells⁴⁶.

R-SMADs (i.e. SMAD2/3) and the co-SMAD SMAD4 are reported to be modulated by lncRNAs. TGF- β /SMAD-induced *TGF- β /SMAD3-interacting long noncoding RNA (lnc-TSI)* binds to the MH2 domain of SMAD3 to diminish its interaction with T β RI in human tubular epithelial cells⁴⁷. *EMT-associated lncRNA induced by TGF β 1 (ELIT-1)* selectively binds to SMAD3, but not SMAD2, and recruits SMAD3 to target gene promoter in multiple cancer cell lines⁴⁸. *ELIT-1* depletion greatly abrogates TGF- β -induced EMT and migration⁴⁸. *LINC00941* functions as a molecular decoy to bind SMAD4 MH2 domain and to protect SMAD4 from being degraded by the E3 ligase β -TrCP in colorectal cancer cell⁴⁹.

Perspectives

The interplay between lncRNAs and TGF- β signaling reveals the important effector role of lncRNAs in TGF- β -induced biological responses and also the intricate and multi-level regulation of TGF- β signaling by lncRNAs to fine-tune its strength and duration. Manipulating critical lncRNA expression in cancer cells may provide a new strategy to target TGF- β -triggered EMT in cancer progression. The tissue-specific expression of lncRNAs can be exploited to selectively target TGF- β signaling in highly-malignant mesenchymal cancer cells to circumvent the on-target effects caused by systemic TGF- β signaling intervention. However, considering the dichotomous role of TGF- β signaling in early and late phases of cancer progression, it is key to understand the mechanisms by which lncRNAs modulate TGF- β signaling in cancer cells in different stages or with difference genetic mutations. Differentially expressed lncRNAs that functionally correlate with TGF- β -induced pro-tumorigenic responses may serve as biomarkers to select cancer patients who can benefit from TGF- β targeted therapies.

References

1. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
2. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101-108 (2012).
3. Volders, P.J. *et al.* An update on LNCipedia: a database for annotated human lncRNA sequences. *Nucleic Acids Res.* **43**, D174-D180 (2015).
4. Uszczynska-Ratajczak, B., Lagarde, J., Frankish, A., Guigo, R. & Johnson, R. Towards a complete map of the human long non-coding RNA transcriptome. *Nat. Rev. Genet.* **19**, 535-548 (2018).
5. Li, Z. *et al.* LncBook 2.0: integrating human long non-coding RNAs with multi-omics annotations. *Nucleic Acids Res.* **51**, D186-D191 (2023).
6. Schlackow, M. *et al.* Distinctive patterns of transcription and RNA processing for human lincRNAs. *Mol. Cell* **65**, 25-38 (2017).
7. Nojima, T. & Proudfoot, N.J. Mechanisms of lncRNA biogenesis as revealed by nascent transcriptomics. *Nat. Rev. Mol. Cell Bio.* **23**, 853-853 (2022).
8. Kutter, C. *et al.* Rapid turnover of long noncoding RNAs and the evolution of gene expression. *Plos Genet.* **8** (2012).
9. Quinn, J.J. *et al.* Rapid evolutionary turnover underlies conserved lncRNA-genome interactions. *Gene Dev.* **30**, 191-207 (2016).
10. Iyer, M.K. *et al.* The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* **47**, 199-208 (2015).
11. Yap, K.L. *et al.* Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* **38**, 662-674 (2010).
12. Rosa, S., Duncan, S. & Dean, C. Mutually exclusive sense-antisense transcription at FLC facilitates environmentally induced gene repression. *Nat. Commun.* **7** (2016).
13. Csorba, T., Questa, J.I., Sun, Q.W. & Dean, C. Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 16160-16165 (2014).

14. Xue, Z.H. *et al.* A G-Rich Motif in the lncRNA Braveheart Interacts with a Zinc-Finger Transcription Factor to Specify the Cardiovascular Lineage. *Mol. Cell* **64**, 37-50 (2016).
15. Zhao, X.Y., Li, S.M., Wang, G.X., Yu, Q. & Lin, J.D. A long noncoding RNA transcriptional regulatory circuit drives thermogenic adipocyte differentiation. *Mol. Cell* **55**, 372-382 (2014).
16. Romero-Barrios, N., Legascue, M.F., Benhamed, M., Ariel, F. & Crespi, M. Splicing regulation by long noncoding RNAs. *Nucleic Acids Res.* **46**, 2169-2184 (2018).
17. Morrissy, A.S., Griffith, M. & Marra, M.A. Extensive relationship between antisense transcription and alternative splicing in the human genome. *Genome Res.* **21**, 1203-1212 (2011).
18. Melo, C.A. *et al.* eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol. Cell* **49**, 524-535 (2013).
19. Grossi, E. *et al.* A lncRNA-SWI/SNF complex crosstalk controls transcriptional activation at specific promoter regions. *Nat. Commun.* **11**, 936 (2020).
20. Xiang, J.F. *et al.* Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res.* **24**, 1150-1150 (2014).
21. Isoda, T. *et al.* Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. *Cell* **171**, 103-119 (2017).
22. Tay, Y., Rinn, J. & Pandolfi, P.P. The multilayered complexity of ceRNA crosstalk and competition. *Nature* **505**, 344-352 (2014).
23. Thomson, D.W. & Dinger, M.E. Endogenous microRNA sponges: evidence and controversy. *Nat. Rev. Genet.* **17**, 272-283 (2016).
24. Wright, B.W., Yi, Z.X., Weissman, J.S. & Chen, J. The dark proteome: translation from noncanonical open reading frames. *Trends Cell Biol.* **32**, 243-258 (2022).
25. Makarewich, C.A. & Olson, E.N. Mining for micropeptides. *Trends Cell Biol.* **27**, 685-696 (2017).
26. Yoon, J.H. *et al.* Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat. Commun.* **4**, 2939 (2013).
27. Sui, B.K. *et al.* A novel antiviral lncRNA, EDAL, shields a T309O-GlcNAcylation site to promote EZH2 lysosomal degradation. *Genome Biol.* **21**, 228 (2020).
28. Andric, V. *et al.* A scaffold lncRNA shapes the mitosis to meiosis switch. *Nat. Commun.* **12**, 770 (2021).
29. Tsai, M.C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689-693 (2010).
30. Massague, J. & Gomis, R.R. The logic of TGF β signaling. *FEBS Lett.* **580**, 2811-2820 (2006).
31. Katsuno, Y., Lamouille, S. & Derynck, R. TGF- β signaling and epithelial-mesenchymal transition in cancer progression. *Curr. Opin. Oncol.* **25**, 76-84 (2013).
32. Richards, E.J. *et al.* Long non-coding RNAs (lncRNA) regulated by transforming growth factor (TGF) β : lncRNA-hit-mediated TGF β -induced epithelial to mesenchymal transition in mammary epithelia. *J. Biol. Chem.* **290**, 6857-6867 (2015).
33. Yuan, J.H. *et al.* A long noncoding RNA activated by TGF- β promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell* **25**, 666-681 (2014).
34. Jiang, L. *et al.* HCP5 is a SMAD3-responsive long non-coding RNA that promotes lung adenocarcinoma metastasis via miR-203/SNAI1 axis. *Theranostics* **9**, 2460-2474 (2019).
35. Terashima, M., Ishimura, A., Wanna-Udom, S. & Suzuki, T. MEG8 long noncoding RNA contributes to epigenetic progression of the epithelial-mesenchymal transition of lung and pancreatic cancer cells. *J. Biol. Chem.* **293**, 18016-18030 (2018).
36. Li, Z., Liu, H., Zhong, Q., Wu, J. & Tang, Z. lncRNA UCA1 is necessary for TGF- β -induced epithelial-mesenchymal transition and stemness via acting as a ceRNA for Slug in glioma cells. *FEBS Open Biol.* **8**, 1855-1865 (2018).
37. Shen, X. *et al.* The long noncoding RNA TUG1 is required for TGF- β /TWIST1/EMT-mediated metastasis in colorectal cancer cells. *Cell Death Dis.* **11**, 65 (2020).
38. Li, R.H. *et al.* Long noncoding RNA ATB promotes the epithelial-mesenchymal transition by upregulating the miR-200c/Twist1 axis and predicts poor prognosis in breast cancer. *Cell Death Dis.* **9**, 1171 (2018).
39. Fan, Y. *et al.* TGF- β -induced upregulation of malat1 promotes bladder cancer metastasis by associating with suz12. *Clin. Cancer Res.* **20**, 1531-1541 (2014).
40. Papoutsoglou, P. *et al.* The TGFB2-AS1 lncRNA regulates TGF- β signaling by modulating corepressor activity. *Cell Rep.* **28**, 3182-3198 (2019).
41. Nakao, A. *et al.* Identification of SMAD7, a TGF- β -inducible antagonist of TGF- β signalling. *Nature* **389**, 631-635 (1997).
42. Kang, Y.B., Chen, C.R. & Massague, J. A self-enabling TGF β response coupled to stress signaling: SMAD engages stress response factor ATF3 for Id1 repression in epithelial cells. *Mol. Cell* **11**, 915-926 (2003).

43. Papoutsoglou, P. *et al.* The noncoding MIR100HG RNA enhances the autocrine function of transforming growth factor β signaling. *Oncogene* **40**, 3748-3765 (2021).
44. Huang, D.L. *et al.* Long noncoding RNA SGO1-AS1 inactivates TGF β signaling by facilitating TGFB1/2 mRNA decay and inhibits gastric carcinoma metastasis. *J. Exp. Clin. Canc. Res.* **40**, 342 (2021).
45. Xu, L. *et al.* Long non-coding RNA SMASR inhibits the EMT by negatively regulating TGF- β /SMAD signaling pathway in lung cancer. *Oncogene* **40**, 3578-3592 (2021).
46. Zhu, L., Liu, Y., Tang, H. & Wang, P. FOXP3 activated-LINC01232 accelerates the stemness of non-small cell lung carcinoma by activating TGF- β signaling pathway and recruiting IGF2BP2 to stabilize TGFB1. *Exp. Cell Res.* **413**, 113024 (2022).
47. Wang, P. *et al.* Long noncoding RNA lnc-TSI inhibits renal fibrogenesis by negatively regulating the TGF- β /SMAD3 pathway. *Sci. Transl. Med.* **10**, eaat2039 (2018).
48. Sakai, S. *et al.* Long noncoding RNA ELIT-1 acts as a SMAD3 cofactor to facilitate TGF β /SMAD signaling and promote epithelial-mesenchymal transition. *Cancer Res.* **79**, 2821-2838 (2019).
49. Wu, N. *et al.* LINC00941 promotes CRC metastasis through preventing SMAD4 protein degradation and activating the TGF- β /SMAD2/3 signaling pathway. *Cell Death Differ.* **28**, 219-232 (2021).

3) Scope of this thesis

In **Chapter 1**, we review the transduction of TGF- β signaling and the intricate regulation of TGF- β signaling at multiple layers. The biphasic role of TGF- β signaling in cancer progression is discussed. We also review the interplay between long non-coding RNAs (lncRNAs) and TGF- β signaling in EMT. In **Chapter 2**, we identified a lncRNA *LITATSI* that functions as a protector of TGF- β -induced EMT in breast and non-small cell lung cancer. *LITATSI* enhances the polyubiquitination and proteasomal degradation of T β RI by strengthening the interaction between T β RI and the E3 ligase SMURF2. *LITATSI* maintains the cytoplasmic localization of SMURF2. In **Chapter 3**, we uncovered an unannotated lncRNA *LETSI* as a novel enforcer of TGF- β signaling and TGF- β -induced EMT in breast and non-small cell lung cancer cells. Mechanistic study revealed that *LETSI* cooperates with NFAT5 to bind *NR4A1* promoter and induces the expression of *NR4A1*, a critical determinant of a destruction complex for inhibitory SMAD7. In **Chapter 4**, we found that a transcriptional repressor Ovo like transcriptional repressor 1 (OVOL1) inhibits TGF- β -induced EMT by facilitating T β RI degradation. We uncovered that OVOL1 interacts with and prevents SMAD7 polyubiquitination and degradation. A small molecule compound 6-formylindolo(3,2-b)carbazole (FICZ) was identified to activate OVOL1 expression and thereby antagonizes (at least in part) TGF- β -mediated EMT and migration in breast cancer cells.



Chapter 2

LncRNA *LITATS1* suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation

Chuannan Fan^{1,2}, Qian Wang^{1,2}, Thomas B Kuipers³, Davy Cats³, Prasanna Vasudevan Iyengar^{1,2}, Sophie C Hagenaars⁴, Wilma E Mesker⁴, Peter Devilee^{5,6}, Rob A E M Tollenaar⁴, Hailiang Mei³, and Peter ten Dijke^{*,1,2}

¹Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands

²Oncode Institute, Leiden University Medical Center, Leiden, The Netherlands

³Department of Biomedical Data Sciences, Sequencing Analysis Support Core, Leiden University Medical Center, Leiden, The Netherlands

⁴Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands

⁵Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands

⁶Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

*Corresponding author. Tel: +31 71 526 9271; E-mail: p.ten_dijke@lumc.nl

Abstract

Epithelial cells acquire mesenchymal phenotypes through epithelial-mesenchymal transition (EMT) during cancer progression. However, how epithelial cells retain their epithelial traits and prevent malignant transformation is not well understood. Here, we report that the long noncoding RNA *LITATSI* (*LINC01137*, *ZC3H12A-DT*) is an epithelial gatekeeper in normal epithelial cells and inhibits EMT in breast and non-small cell lung cancer cells. Transcriptome analysis identified *LITATSI* as a TGF- β target gene. *LITATSI* expression is reduced in lung adenocarcinoma tissues compared with adjacent normal tissues and correlates with a favorable prognosis in breast and non-small cell lung cancer patients. *LITATSI* depletion promotes TGF- β -induced EMT, migration, and extravasation in cancer cells. Unbiased pathway analysis demonstrated that *LITATSI* knockdown potently and selectively potentiates TGF- β /SMAD signaling. Mechanistically, *LITATSI* enhances the polyubiquitination and proteasomal degradation of TGF- β type I receptor (T β RI). *LITATSI* interacts with T β RI and the E3 ligase SMURF2, promoting the cytoplasmic retention of SMURF2. Our findings highlight a protective function of *LITATSI* in epithelial integrity maintenance through the attenuation of TGF- β /SMAD signaling and EMT.

Introduction

At the early stage of epithelium-derived cancers, highly polarized epithelial cells gradually lose cell–cell adhesion and acquire mesenchymal-like features through a process called epithelial-mesenchymal transition (EMT)^{1–3}. This process is characterized by the loss of epithelial markers (E-cadherin, ZO-1, etc.) and the gain of mesenchymal markers (Fibronectin, N-cadherin, Vimentin, etc.) in epithelial cells. Mesenchymal cancer cells can invade through the basement membrane and intravasate into the vascular circulation, resulting in the dissemination of cancer cells and the formation of metastases in distant organs³. However, the reversible EMT process includes multiple intermediate states, referred to as partial or hybrid EMT^{4, 5}. In particular, cancer cells with a dynamic epithelial-mesenchymal plasticity (EMP) phenotype demonstrate greater malignancy, more prominent stem cell characteristics, and greater resistance to chemotherapy^{5–8}.

Signaling by the secreted cytokine transforming growth factor β (TGF- β) is a main EMT driver, and targeting proactive TGF- β signaling for cancer treatment has been evaluated clinically^{9, 10}. TGF- β initiates signaling upon binding to complexes of TGF- β type I and type II serine/threonine receptors (T β RI and T β RII, respectively). Activated T β RI induces regulated (R)-SMAD2/3 phosphorylation, after which phosphorylated SMAD2/3 translocate into the nucleus by forming complexes with SMAD4. These SMAD complexes regulate gene transcription by cooperating with other transcription factors^{11, 12}. The intensity and duration of TGF- β signaling are finely tuned at multiple levels¹³. At the receptor level, SMAD-Specific E3 Ubiquitin Protein Ligase 1/2 (SMURF1/2) are recruited to activated T β RI by interacting with the inhibitory protein SMAD7 and thereby polyubiquitinate and degrade T β RI^{14, 15}.

Long noncoding RNAs (lncRNAs) are defined as transcripts that are longer than 200 nucleotides, transcribed by RNA polymerase II, and lack the protein-coding ability^{16, 17}. The regulatory functions of lncRNAs in various biological processes and pathological events, including cancer progression, have been shown^{18, 19}. LncRNAs can serve as guides, scaffolds or decoys to modulate the interactions between biological macromolecules, such as protein–protein interactions and protein–DNA interactions, and thereby regulate gene expression at multiple levels^{17, 20}. In addition, lncRNAs can sponge microRNAs (miRNAs) by acting as competitive endogenous RNAs (ceRNAs)^{21, 22}.

Epithelial cells protect their integrity by sustaining the expression of epithelial gatekeeper proteins such as OVOL1/2²³, GRHL2²⁴, and C/EBP α ²⁵. Loss of these proteins induces epigenetic reprogramming and/or hyperactivation of EMT-promoting transcription factors or signaling pathways, resulting in the disruption of epithelial integrity and the acquisition of mesenchymal features^{23, 24, 26}. LncRNAs are emerging as a new class of EMT regulators. By functioning as an epigenetic silencer, *human HOX antisense intergenic RNA (HOTAIR)* suppresses EMT and breast cancer metastasis^{27, 28}. However, whether lncRNAs participate in maintaining epithelial architecture is poorly understood. Here, we identify *LncRNA Induced by TGF- β and Antagonizes TGF- β Signaling 1 (LITATSI)* as a protector of epithelial cells to inhibit TGF- β -induced EMT and invasive abilities. Our findings reveal a novel lncRNA-directed mechanism by which epithelial cells maintain their integrity and thereby prevent TGF- β -induced EMT and cancer cell invasion.

Results

***LITATSI* is a cytoplasmic lncRNA whose expression is induced by TGF- β /SMAD signaling.**

TGF- β is a pivotal driver of EMT that disrupts epithelial integrity¹⁰. To investigate the role of lncRNAs in TGF- β -induced EMT and cell migration, we performed transcriptional profiling to screen for TGF- β -induced lncRNAs in breast cell lines that respond to TGF- β -induced EMT (i.e., MCF10A-M1 normal breast epithelial cells and MCF10A-M2 premalignant breast cells; Appendix Fig S1A–C) or in which TGF- β stimulates cell migration and invasion (i.e., MDA-MB-231 mesenchymal triple-negative breast cancer cells; Appendix Fig S1A, B and D). RNA sequencing (RNA-seq) analysis was performed on these three cell lines stimulated with TGF- β for short (2 h), moderate (8 h), and prolonged (24 h) durations (Fig 1A). Using samples without TGF- β treatment (0 h) as the reference, we selected 15 lncRNAs whose expression is decent among the 25 lncRNAs that were induced by TGF- β in at least two of the three cell lines after all TGF- β stimulation durations ($P < 0.05$, fold change > 2 ; Fig 1B, Appendix Fig S1E, Table S1). Analysis of a separate batch of RNA samples from MCF10A-M2 cells validated the induction of these 15 lncRNA hits by TGF- β (Appendix Fig S1F). Moreover, 9 of the 15 lncRNAs were also potently upregulated by TGF- β in A549 lung adenocarcinoma cells, a cell line that is commonly used to investigate TGF- β -induced EMT (Appendix Fig S2A). A further screen directed by individually depleting the 9 lncRNAs with two independent GapmeRs identified two lncRNAs (No. 4 and No. 11; Appendix Table S1) whose knockdown augmented TGF- β -mediated effects on the rearrangement of the actin cytoskeleton into filamentous (F)-actin stress fibers and EMT marker expression (Appendix Fig S2B–D). We observed that one unannotated lncRNA (No. 12; Appendix Table S1) exerted the opposite effects (Appendix Fig S2B–D). As a well-characterized lncRNA, lncRNA No. 4 (*NKILA*) was reported to be induced by TGF- β and alleviate EMT and cancer metastasis^{29–31}. We prioritized lncRNA No. 11 (which we termed *LITATSI*) for further investigation due to its abundant basal and prominent TGF- β -induced expression (Appendix Figs S1F and S2A) and its potent inhibitory effects on TGF- β -induced EMT (Appendix Fig S2C and D).

To evaluate *LITATSI* kinetic expression pattern upon TGF- β treatment, we prolonged the duration of TGF- β stimulation and observed a sustained *LITATSI* expression until 72 h in MDA-MB-231 and A549 cells (Fig EV1A). To verify and extend our identification of *LITATSI* as a TGF- β -induced target gene, we depleted SMAD4 in MDA-MB-231 cells and found that both basal and TGF- β -induced *LITATSI* expression levels were mitigated (Figs 1C and EV1B). Moreover, *LITATSI* expression was enhanced upon ectopic expression of constitutively active TGF- β type I receptor (caT β RI) in HEK293T cells (Fig 1D).

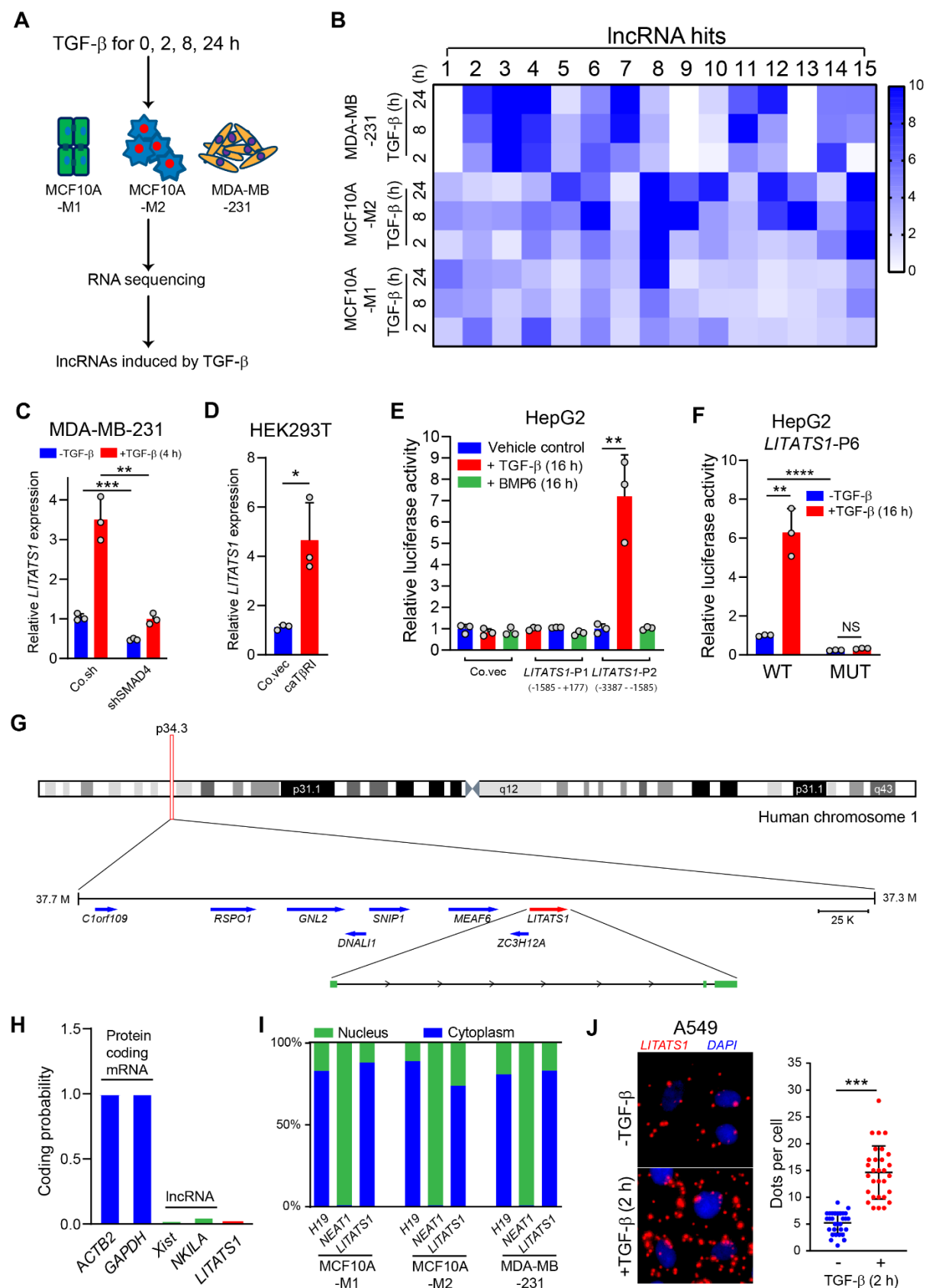


Figure 1. *LITATS1* is a TGF- β -induced lncRNA. (A) Scheme for screening lncRNAs induced by TGF- β . MCF10A-M1, MCF10A-M2, and MDA-MB-231 cells were treated without (0 h) or with TGF- β for 2 h, 8 h, or 24 h. RNA samples (biological triplicates) were collected for RNA-seq, and lncRNAs induced by TGF- β were selected for further analysis. (B) Heatmap showing the log₂ fold changes in the 15 lncRNA hits induced by TGF- β at all three time points (2 h, 8 h, and 24 h vs. 0 h) in at least two cell lines. (C) *LITATS1* expression (as detected by RT-qPCR) upon *SMAD4* knockdown in MDA-MB-231 cells.

LncRNA *LITATS1* suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation

Cells were serum starved for 16 h and TGF- β was added for 4 h. Representative results from a minimum of three independent experiments are shown. (D) *LITATS1* expression (as detected by RT-qPCR) in HEK293T cells. Cells were transfected without (Co.vec) or with the constitutively active TGF- β type I receptor (caT β RI) ectopic expression construct. Representative results from a minimum of three independent experiments are shown. (E) Effect of TGF- β on *LITATS1* promoter activity as determined by luciferase reporter assays. HepG2 cells were transfected with empty pGL4 vector (Co.vec) or with two indicated *LITATS1* promoter luciferase reporters (*LITATS1*-P1 and *LITATS1*-P2). Cells were stimulated with ligand buffer as the vehicle control (-), BMP6 (50 ng/ml), or TGF- β (5 ng/ml) for 16 h. Representative results from a minimum of three independent experiments are shown. (F) Effect of caT β RI and SMAD3 on *LITATS1* promoter activity as determined by luciferase reporter assays. HepG2 cells were transfected with ectopic expression constructs for the *LITATS1* promoter 2 luciferase reporter (*LITATS1*-P2) and caT β RI or SMAD3 and were then stimulated with or without TGF- β for 16 h. Representative results from a minimum of three independent experiments are shown. (G) Schematic representation of the genomic location of *LITATS1* and its neighboring genes. The arrows indicate the direction of transcription. (H) CPAT software was used to predict the coding potential of protein-coding mRNAs (*ACTB2* and *GAPDH*), well-annotated lncRNAs (*Xist* and *NKILA*), and *LITATS1*. (I) Expression analysis of lncRNA *H19*, *NEAT1*, and *LITATS1* expression levels in the cytoplasmic and nuclear fractions of MCF10A-M1, MCF10A-M2, and MDA-MB-231 cells. Representative results from a minimum of three independent experiments are shown. (J) RNA fluorescence in situ hybridization was performed to evaluate *LITATS1* expression and subcellular localization in A549 cells. Cells were treated with or without TGF- β for 2 h. Representative images are shown in the left panel, and signal quantification data are shown in the right panel. Scale bar = 10 μ m. Representative results from two independent experiments are shown.

Data information: TGF- β was applied at a final concentration of 5 ng/ml. (C, D, E, F) are expressed as the mean \pm SD values from three biological replicates (n=3). (J) is expressed as the mean \pm SD values from 30 biological replicates (n=30). *0.01 < P < 0.05; **0.001 < P < 0.01; ***0.0001 < P < 0.001; ****P < 0.0001; NS, not significant. Statistical analysis was based on the unpaired Student's t-test.

To further investigate the mechanism by which TGF- β /SMAD signaling potentiates *LITATS1* expression, the *LITATS1* promoter was characterized. TGF- β but not the closely related family member bone morphogenetic protein (BMP)6, stimulated the transcriptional activity of the *LITATS1* promoter fragment (-3,387 to -1,585 bp upstream of the transcription start site; chromosome 1: 37,476,029 to 37,477,830 (GRCh38.p14)) when placed upstream of a luciferase reporter gene (Figs 1E and EV1C). In addition, ectopic expression of caT β RI or its downstream transcriptional effector SMAD3 (in either the absence or presence of exogenous TGF- β) enhanced *LITATS1* promoter activity (Fig EV1D). Next, transcriptional activity analysis of *LITATS1* promoter truncation mutants demonstrated that the promoter region containing bp -3,212 to -2,649 (chromosome 1: 37,477,093 to 37,477,655 (GRCh38.p14)) was responsible for the TGF- β -mediated transcriptional activity (Fig EV1E). Notably, mutation of a putative SMAD binding site completely abrogated basal and TGF- β -driven *LITATS1* transcription (Figs 1F and EV1F). Collectively, our results reveal that *LITATS1* is a direct target gene of TGF- β /SMAD signaling.

Next, we mapped the *LITATS1* locus on chromosome 1, which is located at head-to-head orientation to a protein-coding gene *ZC3H12A* (Fig 1G). The 5' and 3' rapid amplification of cDNA ends (RACE) assays demonstrated that *LITATS1* is a 1,443 nt three-exon transcript that is identical to an annotated lncRNA *LINC01137* in the NCBI database or *ZC3H12A-DT* in the Ensembl database (Figs 1G and EV1G). Although *LITATS1* is shown as the only splice variant in the NCBI database, *ZC3H12A-DT* was found to be spliced into seven splice variants as shown in the Ensembl database (Appendix Fig S3A). To check whether *LITATS1* (splice variant 1) is the only TGF- β -induced *ZC3H12A-DT* splice variant, we analyzed the RNA-seq data and estimated the raw sequencing reads using StringTie that can discriminate the seven splice variants. We found that *LITATS1* (splice variant 1) basal expression was the highest among the seven splice variants (Appendix Fig S3B). Moreover, *LITATS1* (splice variant 1) was the only variant that can be induced by TGF- β in all three breast cell lines (Appendix Fig S3B). Additionally, reverse transcription-quantitative PCR (RT-qPCR) analysis of MDA-MB-231 and A549 cells consolidated this result (Appendix Fig S3C and D).

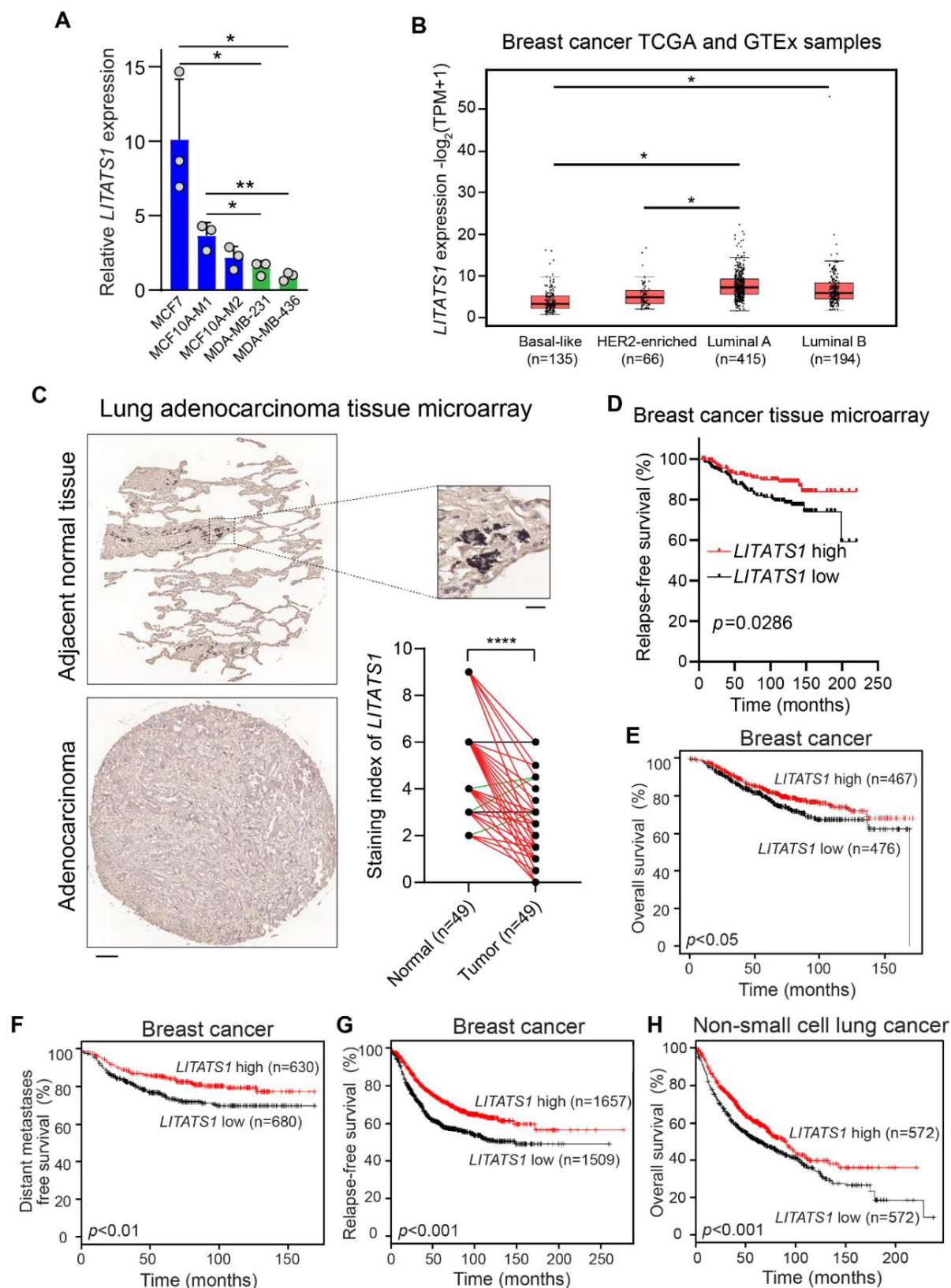


Figure 2. *LITATS1* expression correlates with better prognosis in breast cancer and lung cancer patients. (A) *LITATS1* expression in different breast cells as measured by RT-qPCR. Results from epithelial-like and mesenchymal-like cells are labeled in blue and green, respectively. Representative results from two independent experiments are shown. (B) Comparison of *LITATS1* expression in breast cancer classified by PAM50 subtypes. (C) Quantification of *LITATS1* expression levels by in situ hybridization in lung adenocarcinoma tissue microarrays. Representative images (bar = 100 μm) and zoomed images (bar = 20 μm) of in situ hybridization results in lung adenocarcinoma and matched adjacent normal tissues are shown in the

left panel. The comparison of the *LITATSI* staining index between the paired tissues is shown in the right panel. Tissue pairs with higher *LITATSI* expression in the normal tissue (normal) than in the lung adenocarcinoma tissue (tumor) are highlighted in red, whereas tissue pairs with lower *LITATSI* expression in the normal tissue than in the tumor tissue are highlighted in green. (D) Kaplan–Meier survival curves of relapse-free survival in 175 breast cancer patients stratified by *LITATSI* expression. *LITATSI* expression was measured by in situ hybridization in breast cancer tissue microarrays. (E–H) Kaplan–Meier survival curves of overall survival (E), distant metastasis-free survival (F), and relapse-free survival (G) in breast cancer patients and overall survival (H) in non-small cell lung cancer patients stratified by *LITATSI* expression. The data were generated via Kaplan–Meier Plotter (<https://kmplot.com/analysis/>).

Data information: (A) is expressed as the mean \pm SD values from three biological replicates ($n = 3$). (B) is represented as box-and-whisker plots with 5–95 percentile line representing the median of each group. Numbers below the plot represent patient numbers (biological replicates). (C) is expressed as the mean \pm SD values from 49 biological replicates ($n = 49$). $*0.01 < P < 0.05$; $**0.001 < P < 0.01$; $****P < 0.0001$. In (A, B), statistical analysis was based on the unpaired Student's t-test. In (C), statistical analysis was based on the paired Student's t-test. In (D–H), the log-rank (Mantel-Cox) test was applied to calculate the statistical significance.

Bioinformatic analysis with Coding Potential Assessment Tool (CPAT)³² predicted that *LITATSI* lacked coding potential (Fig 1H). As the subcellular localization of lncRNAs aids in deciphering their functions and mechanisms, subcellular fractionation followed by RT–qPCR was carried out. As shown in Fig 1I, *LITATSI* was localized mainly in the cytoplasm (73.9–88.1%) of three breast cell lines, which was confirmed by fluorescence in situ hybridization in A549 cells (Figs 1J and EV1H). Moreover, TGF- β stimulation did not alter the cytoplasmic and nuclear distribution of *LITATSI* (Fig EV1I). Collectively, these results reveal that *LITATSI* is a cytoplasmic lncRNA whose expression is induced by TGF- β /SMAD signaling.

***LITATSI* expression correlates with a better outcome in cancer patients**

To explore the relationship between *LITATSI* and EMT, *LITATSI* expression was initially analyzed in a panel of breast cell lines with epithelial and/or mesenchymal features. Two mesenchymal-like breast cancer cell lines, MDA-MB-231 and MDA-MB-436, displayed less *LITATSI* expression than three epithelial-like cell lines (MCF10A-M1, MCF10A-M2, and MCF7; Fig 2A). In addition, analysis of RNA-seq data from the TCGA³³ and GTEx³⁴ breast cancer datasets revealed that *LITATSI* expression was significantly decreased in patient samples classified into the basal-like and HER2-enriched subtypes with poor prognosis compared with the luminal A and luminal B subtypes with better prognosis³⁵ (Fig 2B). Moreover, in situ hybridization with a *LITATSI* probe in a commercial tissue microarray showed that *LITATSI* expression was reduced in lung adenocarcinoma samples compared with matched adjacent normal samples, with a lower level in 89.8% (44 of 49) of the tested samples (Fig 2C). To investigate the correlation between *LITATSI* expression and survival in patients with breast cancer, a Kaplan–Meier plot was generated based on the in situ hybridization results in the ORIGO breast cancer tissue microarray³⁶. Higher *LITATSI* expression was associated with a higher relapse-free survival rate ($P = 0.0286$) in the cohort of 175 breast cancer patients (Fig 2D). Furthermore, bioinformatic Kaplan–Meier analysis using other patient cohorts^{37, 38} also revealed that high *LITATSI* expression correlated with a favorable outcome in breast and non-small cell lung cancer patients (Fig 2E–H). Together, our results demonstrate that *LITATSI* is expressed at lower levels in mesenchymal breast cancer cells and that *LITATSI* expression correlates with a favorable clinical outcome in breast and non-small cell lung cancer patients.

Loss of *LITATSI* potentiates TGF- β -induced EMT and cell migration

To further investigate the impact of *LITATSI* on TGF- β -induced EMT, *LITATSI* was overexpressed by CRISPR activation (CRISPRa) in MCF10A-M2 cells (Appendix Fig S4A) or using a lentiviral ectopic expression construct in A549 cells (Appendix Fig S4B). TGF- β -induced the downregulation of E-cadherin expression and the upregulation of mesenchymal marker expression were alleviated upon *LITATSI* ectopic expression in both cell lines (Figs 3A and EV2A).

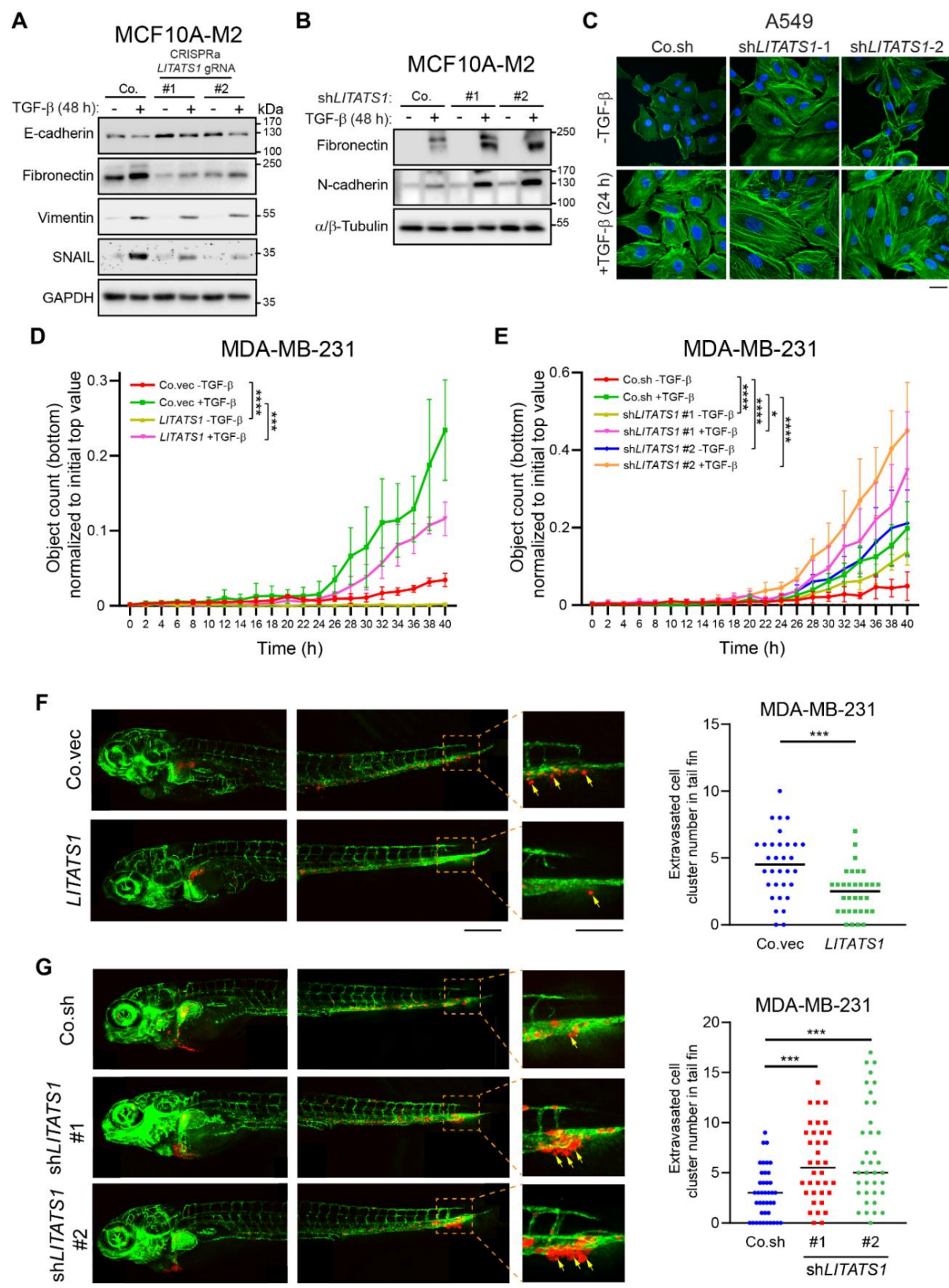


Figure 3. *LITATS1* knockdown potentiates EMT, cell migration, and cell extravasation. (A, B) Effect of *LITATS1* on TGF- β -induced EMT marker expression in MCF10A-M2 upon CRISPRa-mediated *LITATS1* overexpression (A) or shRNA-mediated knockdown (B). GAPDH or α/β -Tubulin, loading control. The results of *LITATS1* overexpression and knockdown are shown in Appendix Fig S4A and Fig EV2B. Representative results from a minimum of three independent experiments are shown. (C) Immunofluorescence analysis of F-actin expression and localization in A549 cells upon shRNA-mediated *LITATS1* depletion. Cells were treated with or without TGF- β for 24 h. Nuclei were visualized by DAPI staining. Scale bar = 30 μ m.

The result of *LITATSI* knockdown is shown in Appendix Fig S4C. Representative results from two independent experiments are shown. (D, E) An IncuCyte chemotactic migration assay was performed to evaluate the effect of *LITATSI* ectopic expression (D) or knockdown (E) on TGF- β -induced MDA-MB-231 cell migration. The results of *LITATSI* overexpression and knockdown are shown in Appendix Fig S4D and E. Representative results from two independent experiments are shown. (F, G) *In vivo* zebrafish extravasation experiments with MDA-MB-231 cells upon ectopic *LITATSI* expression (F) or *LITATSI* knockdown (G). Representative zoomed images of the tail fin area are shown in the left panels. Extravasated breast cancer cell clusters are indicated with yellow arrows. Analysis of the extravasated cell cluster numbers in the indicated groups is shown in the right panels. Whole zebrafish image, bar = 309.4 μ m; zoomed image, bar = 154.7 μ m. Representative results from two independent experiments are shown.

Data information: TGF- β was applied at a final concentration of 1 ng/ml. (D, E) are expressed as the mean \pm SD values from four biological replicates (n=4). (F, G) are expressed as the mean \pm SD values from 30 biological replicates (n=30). *0.01 < P < 0.05; ***0.0001 < P < 0.001; ****P < 0.0001. In (D, E), statistical analysis was based on two-way ANOVA. In (F, G), statistical analysis was based on the unpaired Student's t-test.

On the contrary, *LITATSI* knockdown by two independent shRNA constructs (sh*LITATSI* #1 and #2) enhanced TGF- β -induced expression of two mesenchymal markers in MCF10A-M2 cells (Figs 3B and EV2B). The inhibitory role of *LITATSI* in EMT was also validated in A549 cells via transcriptional profiling and gene set enrichment analysis (GSEA) upon depletion of *LITATSI*. A significant reverse correlation was observed between *LITATSI* expression and a well-established EMT signature (Fig EV2C). Additionally, *LITATSI* depletion facilitated F-actin formation in the absence of TGF- β and further potentiated TGF- β -induced F-actin formation (Fig 3C). Consistent with these results, *LITATSI* ectopic expression suppressed TGF- β -induced cell migration in MDA-MB-231 and A549 cells, as measured by a chemotactic migration assay (Figs 3D and EV2D). By contrast, *LITATSI* depletion in MDA-MB-231 cells augmented TGF- β -induced cell migration (Fig 3E). In agreement with our *in vitro* migration results, the inhibitory effect of *LITATSI* on *in vivo* cell extravasation was observed in a zebrafish embryo breast cancer xenograft model (Figs 3F and G, and EV2E). Taken together, these data indicate that *LITATSI* functions as a critical suppressor of TGF- β -induced EMT and cell migration.

***LITATSI* attenuates TGF- β /SMAD signaling**

Next, we investigated the mechanism by which *LITATSI* affects TGF- β -induced EMT and migration. Given that *ZC3H12A* is a head-to-head neighboring gene of *LITATSI* (Fig 1G), we checked the effect of *LITATSI* misexpression on *ZC3H12A* expression. Of note, *ZC3H12A* mRNA expression remained unchanged upon genetic perturbations of *LITATSI* (Appendix Fig S4A–E). Therefore, to explore the signaling pathways affected by *LITATSI* in an unbiased manner, transcriptome analysis of A549 cells with *LITATSI* depletion was carried out (Appendix Fig S5A). Strikingly, 11 of the 15 genes with the greatest upregulation upon *LITATSI* knockdown were bona fide TGF- β /SMAD target genes (fold change > 1.5, P < 0.05; Appendix Fig S5B). Furthermore, SMAD3 and SMAD4 were among the top enriched transcription factors that contribute to the gene transcription events mediated by *LITATSI* depletion (Appendix Fig S5C). Pathway enrichment analysis showed that TGF- β signaling was the fourth top pathway among the 10 significantly affected pathways by *LITATSI* depletion (Appendix Fig S5D). In addition, GSEA confirmed the positive correlations between *LITATSI* depletion and the TGF- β gene response signature³⁹ (Fig 4A). Next, we evaluated the effect of *LITATSI* on TGF- β /SMAD signal transduction using a highly selective synthetic SMAD3/4-driven transcriptional reporter⁴⁰. *LITATSI* overexpression suppressed, but *LITATSI* depletion potentiated the TGF- β /SMAD3/4-induced transcriptional response in HepG2 cells (Figs 4B and EV3A). Notably, a *LITATSI* mutant (MUT) in which all the putative start codons were mutated (ATG to ATT) exhibited an inhibitory effect on the TGF- β /SMAD3/4-induced transcriptional response similar to that of wild-type (WT) *LITATSI* (Fig EV3B). This finding was consistent with the prediction of *LITATSI* to be a lncRNA that does not encode small

peptides despite its cytoplasmic localization.

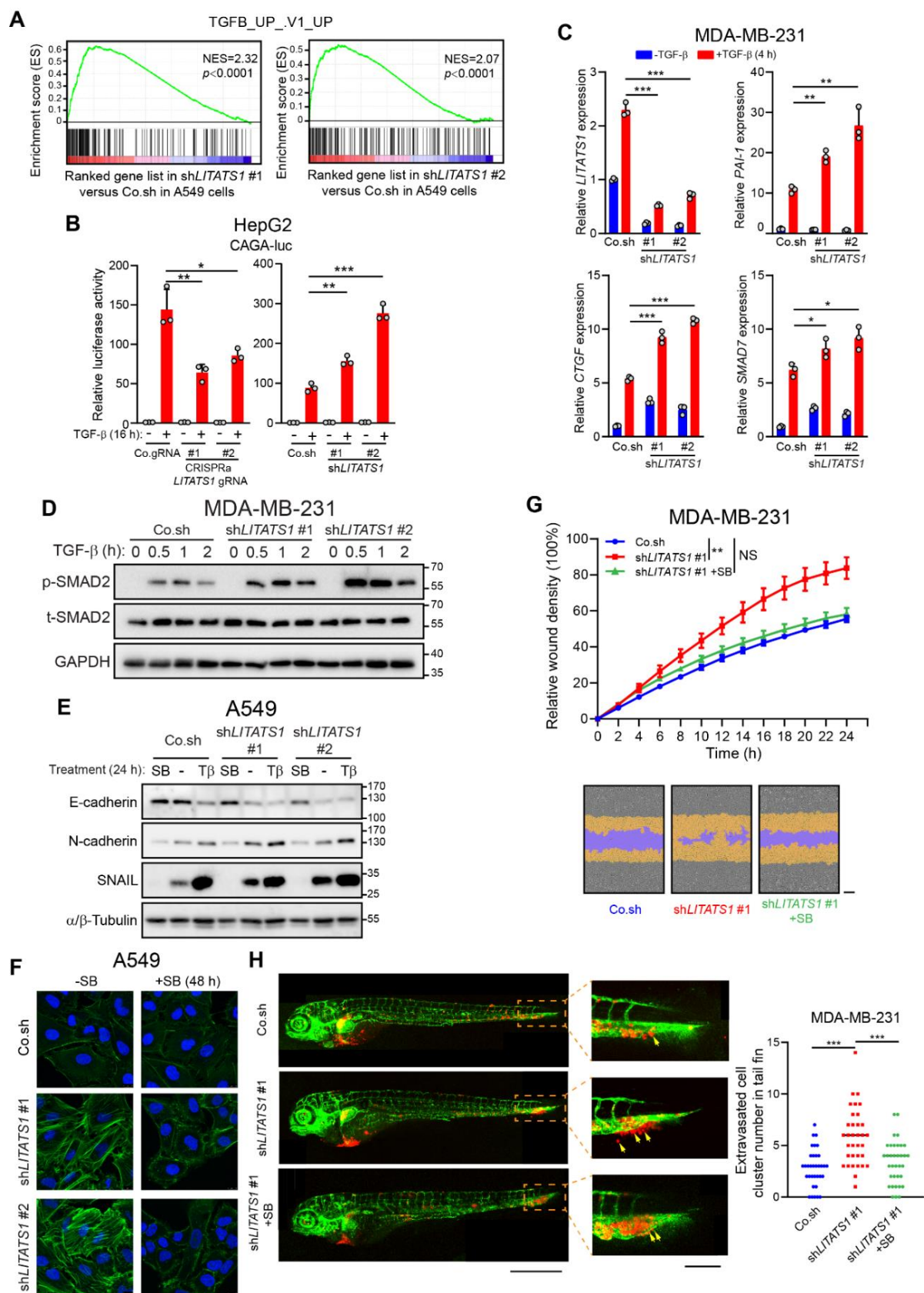


Figure 4. *LITATS1* suppresses TGF- β /SMAD signaling and EMT. (A) GSEA of positive correlations between (manipulated) *LITATS1* expression and the TGF- β gene response signature. (B) Effect of *LITATS1* misexpression on TGF- β /SMAD3

transcriptional activity in HepG2 cells. Cells were transfected with expression constructs for the TGF- β -induced SMAD3/4-dependent CAGA-luc transcriptional reporter and *LITATS1* misexpression. The results of *LITATS1* misexpression are shown in Fig EV3A. Representative results from a minimum of three independent experiments are shown. (C) Expression of TGF- β target genes (as measured by RT-qPCR) in MDA-MB-231 cells without (Co.sh) or with (sh#1 and sh#2) *LITATS1* depletion. Cells were serum starved for 16 h and treated with or without TGF- β for 4 h. Representative results from a minimum of three independent experiments are shown. (D) Effect of *LITATS1* knockdown on TGF- β -induced SMAD2 phosphorylation in MDA-MB-231 cells. Cells were serum starved for 16 h and stimulated with TGF- β for the indicated durations. The p-SMAD2 and total SMAD2 (t-SMAD2) levels were analyzed by western blotting. GAPDH, loading control. Representative results from a minimum of three independent experiments are shown. (E) Effect of *LITATS1* knockdown on E-cadherin, N-cadherin, and SNAIL expression in A549 cells. Cells were stimulated with vehicle control (-), SB431542 (SB; 10 μ M), or TGF- β (T β) for 24 h, and protein expression was analyzed by western blotting. α / β -Tubulin, loading control. Representative results from a minimum of three independent experiments are shown. (F) Effect of *LITATS1* depletion (using two independent shRNAs, i.e., sh*LITATS1* #1 and #2) on F-actin expression and localization (as evaluated by immunofluorescence) in A549 cells. DAPI staining was performed to visualize nuclei. Cells were stimulated with or without SB431542 (SB; 10 μ M) for 48 h. Scale bar = 30 μ m. Representative results from two independent experiments are shown. (G) IncuCyte wound healing migration assays were performed to evaluate the effect of TGF- β signaling inactivation on MDA-MB-231 cell migration mediated by *LITATS1* knockdown. Cells were treated with or without SB431542 (SB; 10 μ M) during the migration assays. Representative results from two independent experiments are shown. (H) *In vivo* zebrafish extravasation experiments with MDA-MB-231 cells upon *LITATS1* knockdown and blockage of TGF- β signaling. Representative zoomed images of the tail fin area are shown in the left panels. Extravasated breast cancer cell clusters are indicated with yellow arrows. Analysis of the extravasated cell cluster numbers in the indicated groups is shown in the right panel. Whole zebrafish image, bar = 618.8 μ m; zoomed image, scale bar = 154.7 μ m. Representative results from two independent experiments are shown.

Data information: TGF- β was applied at a final concentration of 1 ng/ml. (B, C) are expressed as the mean \pm SD values from three biological replicates (n = 3). (G) is expressed as the mean \pm SD from seven biological replicates (n = 7). (H) is expressed as the mean \pm SD values from 30 biological replicates (n = 30). *0.01 < P < 0.05; **0.001 < P < 0.01; ***0.0001 < P < 0.001; NS, not significant. In (B, C, H), statistical analysis was based on the unpaired Student's t-test. In (G), statistical analysis was based on two-way ANOVA followed by Tukey's multiple comparisons test.

Moreover, *LITATS1* knockdown promoted the expression of TGF- β /SMAD target genes in MDA-MB-231 and MCF10A-M2 cells (Figs 4C and EV3C). By contrast, *LITATS1* overexpression attenuated TGF- β /SMAD-induced target gene expression in both cell lines (Fig EV3D and E). Furthermore, TGF- β -induced SMAD2 phosphorylation, which is an immediate downstream indicator of T β RI activity, was promoted in MDA-MB-231 and MCF10A-M2 cells with *LITATS1* depletion (Figs 4D and EV3F). However, TGF- β -induced SMAD2 phosphorylation was mitigated upon ectopic *LITATS1* expression in MDA-MB-231 and MCF10A-M2 cells (Fig EV3G-I). Moreover, the negative regulatory effect of *LITATS1* on TGF- β /SMAD signaling was confirmed by *LITATS1* misexpression in A549 cells (Appendix Fig S5E-G).

We then determined whether the effect of *LITATS1* on uncontrolled EMT is dependent on TGF- β signaling regulation. The *LITATS1* depletion-mediated changes in EMT marker expression and F-actin formation were mitigated by blockade of TGF- β /SMAD signaling with the selective T β RI kinase inhibitor SB431542 in A549 cells (Fig 4E and F, Appendix Fig S5H). Moreover, SB431542 treatment blocked the migration of MDA-MB-231, MCF10A-M2, and A549 cells and the *in vivo* extravasation of MDA-MB-231 cells that were induced by *LITATS1* knockdown (Figs 4G and H, and EV3J, Appendix Fig S5I). Taken together, these results indicate that TGF- β receptor signaling activation is pivotal for the promoting effects on EMT, cell migration, and extravasation that occur upon *LITATS1* depletion.

***LITATS1* destabilizes T β RI by potentiating its polyubiquitination**

The promotion of T β RI-induced SMAD2 phosphorylation resulting from the absence of *LITATS1* (Fig 4D) prompted us to check whether *LITATS1* affects the expression of its upstream TGF- β receptor. We found that upon *LITATS1* ectopic expression, MDA-MB-231 and MCF10A-M2 cells exhibited less T β RI protein expression (Figs 5A and EV4A). This was further confirmed by ectopic expression of *LITATS1* in caT β RI-overexpressing HEK293T cells (Fig EV4B). Interestingly, *TBRI* mRNA expression remained unaffected (Fig EV4A and C).

Consistent with these results, depletion of *LITATS1* enhanced T β RI expression at the protein but not at the mRNA level (Figs 5B and EV4C and D). These results suggest that *LITATS1* may alter T β RI protein turnover. Consistent with this idea, *LITATS1* exerted a negative effect on T β RI protein stability, as measured by a cycloheximide (CHX)-directed time-course assay (Figs 5C and D, and EV4E). To decipher whether lysosomes or proteasomes play a role in the inhibitory effect of *LITATS1* on T β RI protein stability, *LITATS1*-overexpressing MDA-MB-231 and HEK293T cells were challenged with selective chemical lysosome or proteasome inhibitors. *LITATS1*-induced T β RI downregulation was restored only by the proteasome inhibitor MG132 but not by either of the two tested lysosome inhibitors (bafilomycin A1 (BafA1) and hydroxychloroquine (HCQ); Figs 5E and EV4F). Consistent with these results, ectopic *LITATS1* expression greatly increased the T β RI polyubiquitination level (Fig 5F). In addition, *TBR1* knockdown alleviated the induction of EMT resulting from *LITATS1* knockdown in A549 cells (Fig EV4G), suggesting that T β RI is an indispensable target of *LITATS1* in its regulation of EMT. Taken together, these results indicate that *LITATS1* potentiates T β RI polyubiquitination and degradation.

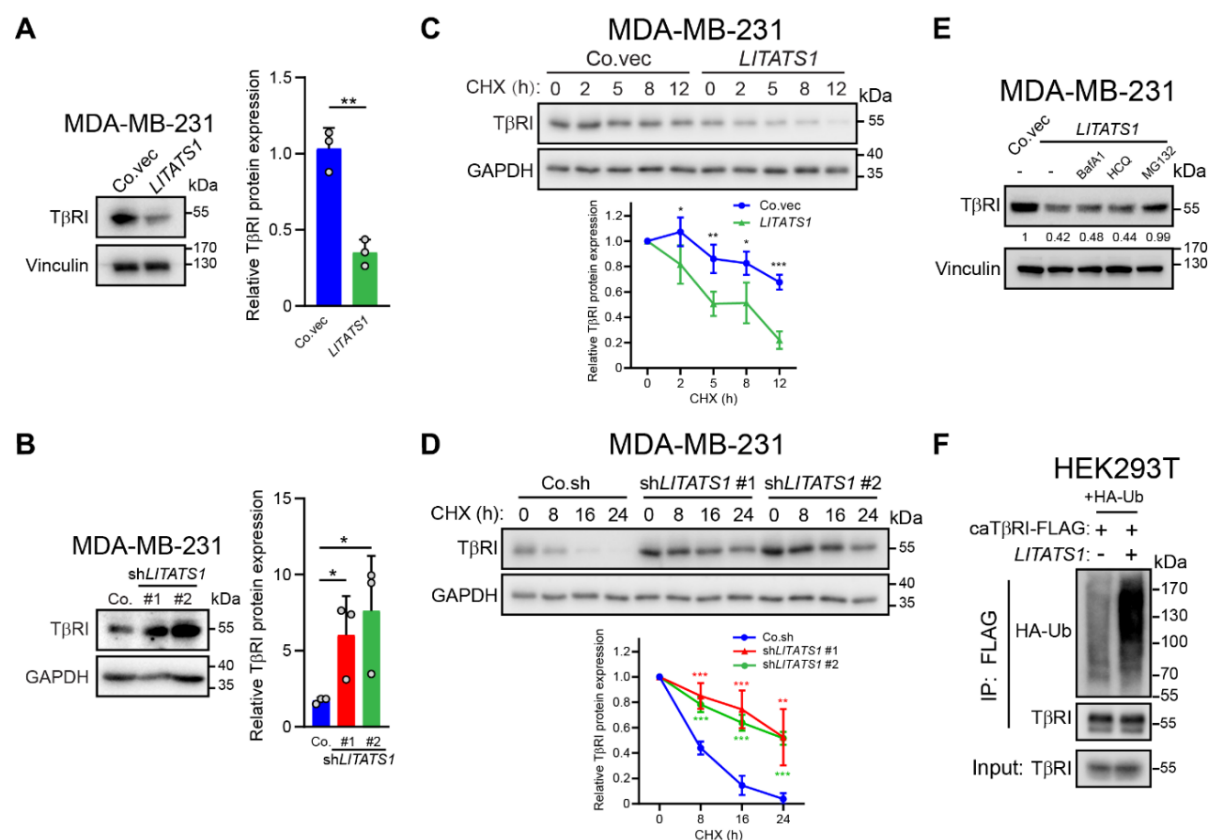


Figure 5. *LITATS1* promotes the polyubiquitination and degradation of T β RI. (A, B) Effect of ectopic *LITATS1* expression (A) or *LITATS1* knockdown (B) on T β RI expression in MDA-MB-231 cells. Right panel: quantification of relative T β RI protein levels. Vinculin or GAPDH, loading control. Representative blots from a minimum of three independent experiments are shown. (C, D) Analysis of T β RI protein stability (as measured by western blotting) in MDA-MB-231 cells with ectopic *LITATS1* expression (C) or *LITATS1* knockdown (D). Cells were treated with CHX (50 μ g/ml) for the indicated durations. Quantification of the relative T β RI protein level is shown in the lower panels. GAPDH, loading control. Representative blots from a minimum of three independent experiments are shown. (E) T β RI expression in MDA-MB-231 cells with ectopic *LITATS1* expression in the absence or presence of lysosome or proteasome inhibitors. Cells were incubated with vehicle control DMSO (-), the lysosome inhibitor BafA1 (20 nM) or HCQ (20 μ M), or the proteasome inhibitor MG132 (5 μ M) for 8 h. Vinculin, loading control. Representative results from a minimum of three independent experiments are shown. (F) Effect of *LITATS1* on T β RI polyubiquitination. HEK293T cells were transfected with ectopic expression constructs for HA-Ubiquitin (HA-Ub), caT β RI-FLAG, and/or *LITATS1*. T β RI polyubiquitination was analyzed by western blotting. Representative blots from a minimum of three independent experiments are shown.

Data information: (A, B) are expressed as the mean \pm SD values from three biological replicates (n = 3). (C, D) are expressed as the mean \pm SD values from four biological replicates (n = 4). *0.01 < P < 0.05; **0.001 < P < 0.01; ***0.0001 < P < 0.001. Statistical analysis was based on the unpaired Student's t-test.

***LITATS1* interacts with T β RI and SMURF2**

To reveal the mechanism by which *LITATS1* increases T β RI polyubiquitination, RNA immunoprecipitation (RIP) coupled with RT-qPCR was performed on lysates from HEK293T cells with ectopic expression of different FLAG-tagged TGF- β /SMAD signaling components (i.e., SMAD2, SMAD3, SMAD4, and T β RI) or modulators (i.e., SMAD7 and the E3 ubiquitin ligases SMURF1/2). Notably, only caT β RI and SMURF2 but not the other ectopically expressed proteins, were capable of coprecipitating *LITATS1* (Figs 6A and EV5A). A recently developed CRISPR-assisted RNA-protein interaction detection method (CARPID)⁴¹, which incorporates CRISPR-CasRx-mediated RNA targeting and proximity labeling to verify endogenous interactions between lncRNAs and proteins of interest, was utilized to validate the interaction between *LITATS1* and T β RI or SMURF2. The specific sequence-matching gRNAs can direct the TurboID-dCasRx complex to *LITATS1*, where RNA-binding proteins in close proximity to *LITATS1* can be labeled with biotin and analyzed by western blotting after enrichment with streptavidin beads (Fig EV5B). We selected the two most effective gRNAs (gRNA #1 and #2) based on the CasRx-directed degradation of *LITATS1* as measured by RT-qPCR (Fig EV5C). As expected, overexpression of these two independent gRNAs increased the biotinylation level of SMURF2 and T β RI in MDA-MB-231 cells (Fig 6B). This effect was further enhanced in the presence of TGF- β and was likely mediated by TGF- β -induced *LITATS1* expression. We further confirmed the interactions between *LITATS1* and T β RI or SMURF2 at the endogenous level in MDA-MB-231 cells using RIP analysis (Fig 6C). When we checked TGF- β -induced *LITATS1* expression, we observed that TGF- β stimulation could not further enhance the interactions between *LITATS1* and T β RI or SMURF2 (Figs EV5D and 5E). To orthogonally confirm these findings, we performed RNA pull-down assays using biotinylated *LITATS1* and negative controls, including biotinylated antisense *LITATS1* and 25 \times poly(A), and proteins produced from HEK293T cells. Western blot analysis revealed that *LITATS1* bound to caT β RI and SMURF2 proteins (Fig 6D). *In vitro* RIP analysis using *in vitro*-transcribed *LITATS1* and the recombinant protein of T β RI intracellular domain (ICD) further confirmed the direct interaction between *LITATS1* and T β RI (Appendix Fig S6A). Moreover, *in vitro* RNA pull-down showed that recombinant SMURF2 but not its homologous protein SMURF1 could coprecipitate with *LITATS1* (Fig 6E). Given that SMURF2 can be recruited to T β RI with the aid of SMAD7, thereby promoting T β RI polyubiquitination and degradation¹⁴, we reasoned that *LITATS1* may also serve as a scaffold to potentiate SMURF2-T β RI interaction. The results of the proximity ligation assay (PLA) in A549 cells demonstrated that TGF- β stimulation resulted in a three-fold increase of SMURF2-T β RI interaction, which was mitigated upon *LITATS1* knockdown (Figs 6F and EV5F). Of note, we found a moderate induction of *LITATS1* (1.7-fold increase) upon 2 h TGF- β stimulation in the RNA-seq analysis of A549 cells, indicating that *LITATS1* promotes T β RI-SMUR2 interaction also independent from its induction by TGF- β , likely by acting as a scaffold. Furthermore, ectopic *LITATS1* expression enhanced SMURF2-induced T β RI polyubiquitination (Fig 6G). Importantly, SMURF2 depletion markedly diminished the increase in T β RI polyubiquitination induced by ectopic *LITATS1* expression (Fig 6H). This latter result indicates that SMURF2 is a key E3 ubiquitin ligase partner of *LITATS1* by which it mediates T β RI polyubiquitination.

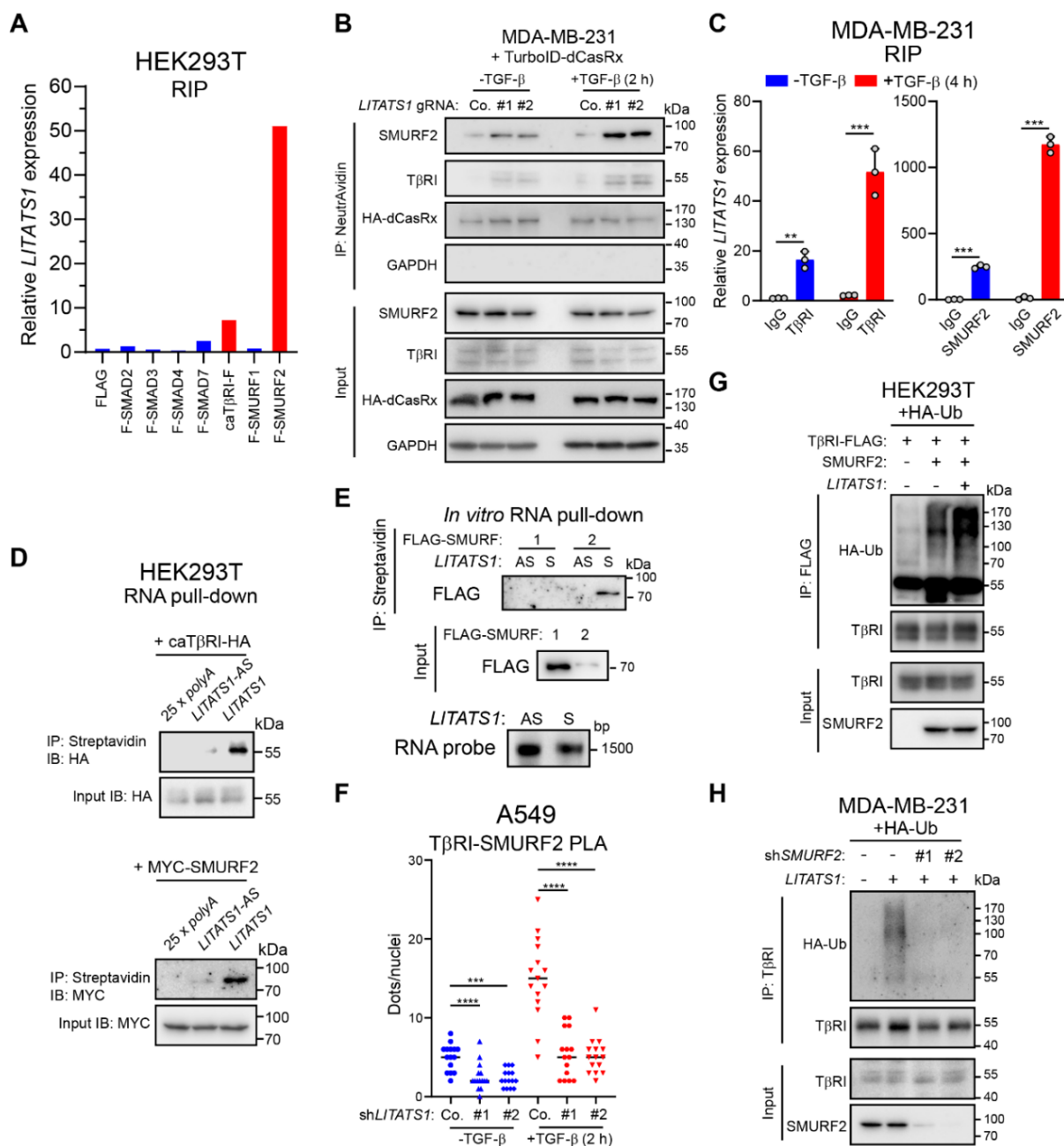


Figure 6. *LITATS1* interacts with TβRI and SMURF2. (A) Interactions between *LITATS1* and TGF-β/SMAD signaling components or modulators were analyzed by RIP. RT-qPCR was performed to detect *LITATS1* expression in immunoprecipitates from HEK293T cells transfected with expression constructs for the indicated proteins. Representative results from two independent experiments are shown. (B) Interactions between *LITATS1* and TβRI or SMURF2 in MDA-MB-231 cells were detected by the CARPID approach. Cells with stable expression of TurboID-dCasRx were transduced without (Co.) or with (#1 and #2) *LITATS1* targeting gRNAs. Cells were stimulated with or without TGF-β (2.5 ng/ml) for 2 h and were then stimulated with biotin (500 μM) for 30 min. Western blotting was performed to detect SMURF2 and TβRI expression in whole-cell lysates (Input) and immunoprecipitates (IP). GAPDH and HA-dCasRx expression levels were measured for equal loading of input samples and as the negative control or positive control, respectively, for proximity biotinylation in immunoprecipitate (IP) samples. Representative results from a minimum of three independent experiments are shown. (C) Interactions between *LITATS1* and TβRI (left) or SMURF2 (right) were analyzed by RIP. MDA-MB-231 cells were stimulated with or without TGF-β (5 ng/ml) for 4 h before RIP. RT-qPCR was performed to detect *LITATS1* expression in immunoprecipitates from MDA-MB-231 cells. IgG was included as the control for immunoprecipitation. Representative results from two independent experiments are shown. (D) Interactions between *LITATS1* and caTβRI or SMURF2 were analyzed by RNA pull-down. Biotinylated 25x poly(A), antisense *LITATS1* (*LITATS1*-AS), or *LITATS1* was incubated with lysates from HEK293T cells transfected with the caTβRI-HA or MYC-SMURF2 expression construct. Western blot analysis was performed to detect HA or MYC expression in whole-cell lysates (Input) and immunoprecipitates (IP). Representative blots from a minimum of three independent experiments are shown. (E) *In vitro* RNA pull-down assays were performed to evaluate the interactions between *LITATS1* and SMURF1/2. *In vitro*-transcribed antisense *LITATS1* (*LITATS1*-AS) or *LITATS1*

(*LITATS1-S*) was incubated with recombinant FLAG-tagged SMURF1 or SMURF2 protein. Western blotting analysis was performed to evaluate FLAG expression in input and IP samples. The amounts of RNA probes used for RNA pull-down were evaluated by agarose gel electrophoresis. Representative results from a minimum of three independent experiments are shown. (F) Quantification of T β RI-SMURF2 PLA in A549 cells with or without *LITATS1* knockdown were treated with or without TGF- β (5 ng/ml) for 2 h. Representative images are shown in Fig EV5F. (G) Effect of *LITATS1* overexpression on SMURF2-mediated T β RI polyubiquitination. HEK293T cells were transfected with expression constructs for HA-Ubiquitin (HA-Ub) and caT β RI-FLAG and ectopic expression constructs for SMURF2 and/or *LITATS1*. Polyubiquitination of T β RI was evaluated by western blotting. Representative blots from a minimum of three independent experiments are shown. (H) Effect of SMURF2 knockdown on *LITATS1*-mediated T β RI polyubiquitination. MDA-MB-231 cells with stable HA-Ub expression were transduced with expression constructs for *LITATS1* and/or two different *SMURF2* shRNAs, as indicated. Polyubiquitination of T β RI was evaluated by western blotting. Representative blots from a minimum of three independent experiments are shown.

Data information: (C) is expressed as the mean \pm SD values from three (n = 3) biological replicates. (F) is expressed as the mean values from 15 (n = 15) biological replicates. **0.001 < P < 0.01; ***0.0001 < P < 0.001; ****P < 0.0001. Statistical analysis was based on the unpaired Student's t-test.

***LITATS1* binds to the WW1 domain of SMURF2 and promotes the cytoplasmic retention of SMURF2**

To map the SMURF2 binding region in *LITATS1*, interactions between SMURF2 and full-length *LITATS1* (Appendix Fig S6B) or four RNA fragments, each representing approximately one-fourth of the *LITATS1* sequence, were evaluated by RNA pull-down (Appendix Fig S6C–F). We observed an interaction between SMURF2 and only the *LITATS1* 5' fragment (T1; 1–350 nt), although this binding was impaired compared with that between SMURF2 and full-length *LITATS1* (Fig 7A). Moreover, analysis of the binding capability of *LITATS1* to SMURF2 truncation mutants demonstrated that the WW1 domain, which is not present in SMURF1, was essential for the binding of SMURF2 to *LITATS1* (Fig 7B–D).

As SMURF2 is translocated from the nucleus to the cytoplasm in response to TGF- β ¹⁴ via a not well-characterized mechanism, we next investigated whether cytoplasmic *LITATS1* alters the subcellular distribution of SMURF2. The immunofluorescence results revealed that *LITATS1* depletion decreased the proportion of cytoplasmic SMURF2 in A549 cells (Fig 7E). Additionally, subcellular fractionation confirmed the attenuation of SMURF2 cytoplasmic localization upon the loss of *LITATS1* (Fig 7F). Consistent with these results, more SMURF2 was retained in the cytoplasm in *LITATS1*-overexpressing MDA-MB-231 cells (Appendix Fig S6G). Of note, SMURF2 protein expression was not affected upon *LITATS1* knockdown (Appendix Fig S6H). Taken together, these results indicate that *LITATS1* potentiates the cytoplasmic retention of SMURF2 without affecting its expression.

Discussion

In this study, we identified *LITATS1* as a critical determinant of epithelial integrity maintenance and inhibitor of TGF- β -induced EMT in breast and non-small cell lung cancer cells. *LITATS1* suppresses TGF- β /SMAD signaling by interacting with SMURF2 and T β RI and promoting the cytoplasmic retention of SMURF2. T β RI polyubiquitination and proteasomal degradation are potentiated by *LITATS1*, leading to suppression of TGF- β /SMAD signaling, TGF- β -induced EMT, and cell migration/invasion (Fig 7G).

We showed that *LITATS1* is the only *ZC3H12A-DT* splice variant that can be induced by TGF- β in breast and non-small cell lung cancer cell lines. To further exclude the involvement of the other six splice variants in the *LITATS1*-mediated effects, we checked the sequences of *LITATS1*-targeting shRNAs and GapmeRs (Appendix Fig S3A). ShRNA #2 and two GapmeRs target the exon 4 that is shared by variants 1, 4, 5, and 6, while shRNA #1 can specifically target the exon 3, which exists only in *LITATS1*. Consistent with the results that *LITATS1* is the main TGF- β -induced *ZC3H12A-DT* splice variant, shRNA #1-mediated *LITATS1* knockdown

affected TGF- β -induced EMT as potent as shRNA #2 and two GapmeRs (Fig 3B and C, and Appendix Fig S2C and D). In addition, the effects of both shRNA constructs on TGF- β signaling regulation are similar (Figs 4 and 5). These results suggest that *LITATS1* is the only *ZC3H12A-DT* splice variant that plays a role in our study.

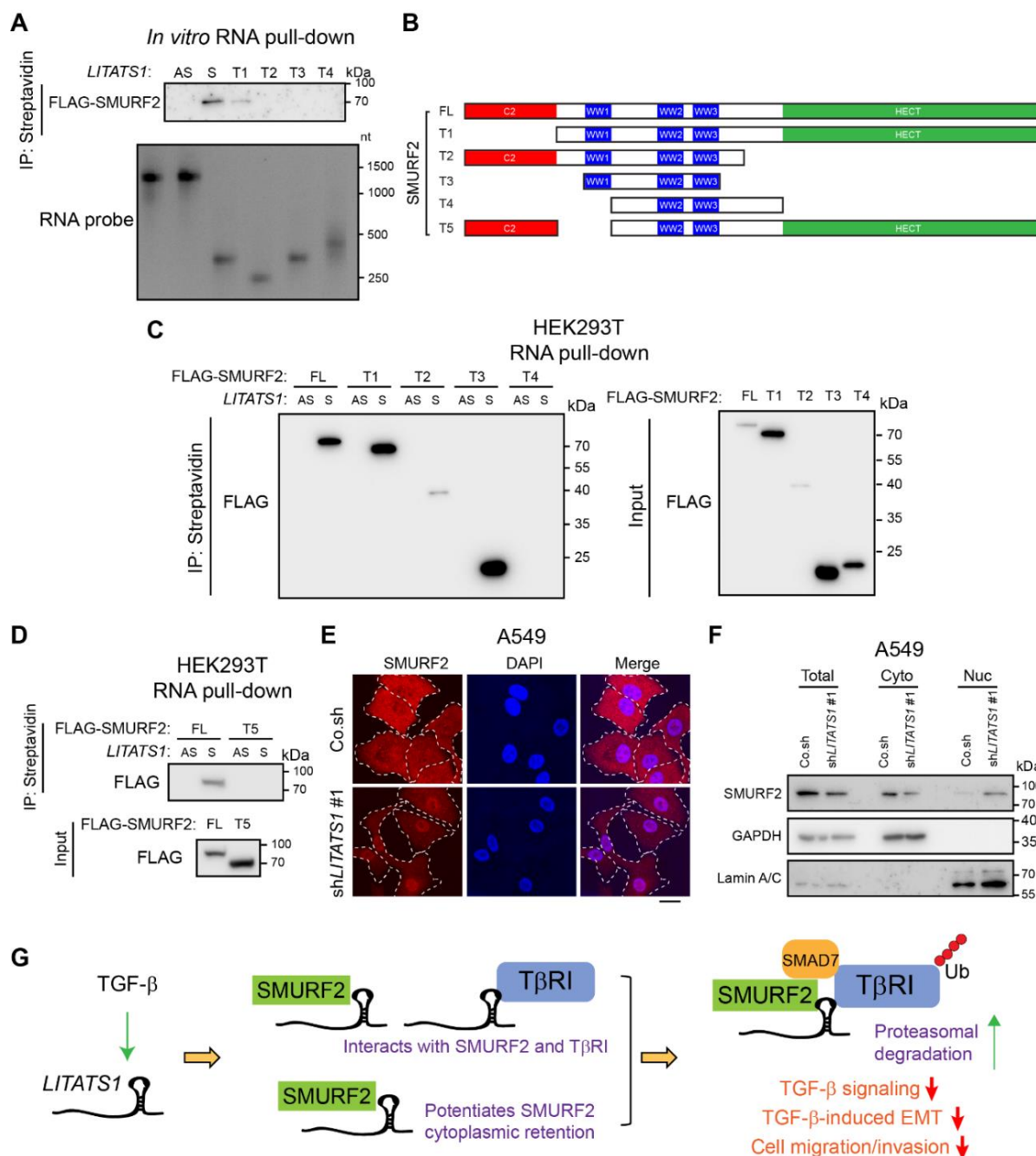


Fig. 7 *LITATS1* retains SMURF2 in the cytoplasm. (A) An *in vitro* RNA pull-down assay was performed to evaluate the interaction between *LITATS1* truncation mutants and SMURF2. Recombinant FLAG-SMURF2 protein was incubated with antisense *LITATS1* (*LITATS1-AS*), *LITATS1* (*LITATS1-S*), or *LITATS1* truncation mutants (T1-T4). Western blot analysis was performed to evaluate FLAG expression in immunoprecipitates (IP). The amounts of RNA probes used for RNA pull-down were evaluated by agarose gel electrophoresis. Representative results from a minimum of three independent experiments are shown. (B) Schematic representation of full-length SMURF2 (FL) and the truncation mutants (T1-T5) tested. (C, D) RNA pull-down assays were performed to evaluate the interaction between *LITATS1* and full-length SMURF2 or its truncation mutants (T1-T5) expressed in HEK293T cells. Western blotting analysis was performed to evaluate FLAG expression in input and immunoprecipitate (IP) samples. Representative results from a minimum of three independent experiments are shown. (E) SMURF2 expression and localization (as measured by immunofluorescence) upon *LITATS1* depletion in A549 cells. DAPI staining was performed to visualize nuclei. Scale bar = 23.2 μ m. Representative results from two independent experiments are

shown. (F) Effect of *LITATS1* knockdown on SMURF2 localization in A549 cells. After subcellular protein fractionation, western blotting was performed to detect SMURF2 expression in whole-cell lysates (Total) and the cytoplasmic (Cyto) and nuclear (Nuc) fractions. The levels of the cytoplasmic marker GAPDH and the nuclear marker Lamin A/C are included to demonstrate subcellular protein fractionation. Representative results from two independent experiments are shown. (G) Schematic working model. TGF- β -induced *LITATS1* interacts with T β RI and SMURF2 and potentiates cytoplasmic retention of SMURF2. *LITATS1* potentiates T β RI polyubiquitination and proteasomal degradation, resulting in suppression of TGF- β signaling, TGF- β -induced EMT, and cancer cell migration/invasion.

TGF- β /SMAD-induced *LITATS1* mitigates T β RI protein turnover and thereby suppresses TGF- β /SMAD signal transduction. Frequently, the products of genes that are transcriptionally induced by TGF- β act in negative or positive feedback loops to fine-tune the intensity and/or duration of TGF- β signaling responses^{42, 43} or participate as effectors in TGF- β -induced biological impacts^{44, 45}. These scenarios also apply to lncRNAs. TGF- β signaling can induce the expression of multiple lncRNAs, e.g., *lncRNA-ATB* and *lncRNA-HIT*, which function as effectors of TGF- β -induced responses^{46, 47}. In addition, certain lncRNAs can act as modulators of TGF- β signaling by altering the expression or activity of TGF- β signaling components⁴⁸⁻⁵². Several lines of evidence indicate that *TBRI* mRNA expression is regulated by lncRNAs at both the transcriptional⁵² and post-transcriptional⁵³⁻⁶¹ levels. However, our results reveal a novel mechanism by which T β RI protein stability is modulated through lncRNA-mediated post-translational modification. LncRNAs have been reported to modulate protein polyubiquitination. *Vimentin-associated lncRNA (VAL)* binds to Vimentin and abrogates Trim16-mediated Vimentin polyubiquitination⁶². In senescent cells, *HOTAIR* facilitates the polyubiquitination of Ataxin-1 and Dzip3 by promoting their associations with the E3 ubiquitin ligases Snurportin-1 and Mex3b, respectively⁶³. For *LITATS1*, SMURF2 appears to be necessary to potentiate T β RI polyubiquitination. However, the contributions of other E3 ubiquitin ligases to this process cannot be excluded.

We mapped the binding region of SMURF2 in the 5' fragment of *LITATS1* (*LITATS1-T1*) by analyzing the *LITATS1* truncation mutants. However, SMURF2 could not interact as potently with *LITATS1-T1* as full-length *LITATS1*. Considering the importance of lncRNA folding structure for its interactions with proteins^{64, 65}, it is highly possible that *LITATS1* truncation may impair its original folding structure that is required for SMURF2 binding. Therefore, checking interactions between SMURF2 and *LITATS1* mutants with small deletions or nucleotide substitutions that elicit minimal effects on *LITATS1* folding can better explore the SMURF2 binding region in *LITATS1*. Moreover, other approaches such as cross-linking and immunoprecipitation (CLIP) coupled with RNA footprinting⁶⁶ can be applied to identify the binding sites of SMURF2 or T β RI in *LITATS1* in live cells.

In response to TGF- β stimulation, SMAD7 binds SMURF2 to activate the ubiquitin ligase activity of SMURF2 by suppressing its autoinhibition and recruits SMURF2 to target T β RI for degradation. Similar to SMAD7, *LITATS1* may serve as a scaffold to facilitate the T β RI-SMURF2 interaction. However, we found that *SMAD7* knockdown mitigated *LITATS1*-directed polyubiquitination of T β RI (Appendix Fig S6I), demonstrating that SMAD7 is required for *LITATS1* to exert its effect on T β RI. Our RIP results suggested a weak interaction between SMAD7 and *LITATS1* (Fig 6A) that is less potent than the interactions between *LITATS1* and T β RI/SMURF2. These results can be explained by the possibility that *LITATS1* is a component of the T β RI/SMURF2/SMAD7 complex and therefore indirectly binds SMAD7. However, SMAD7 knockdown does not affect the interaction between *LITATS1* and T β RI, indicating that the binding of *LITATS1* to T β RI is SMAD7-independent (Appendix Fig S6J). Therefore, further study is required to investigate whether SMAD7 and *LITATS1* function in an additive manner to facilitate the SMURF2/T β RI complex formation.

We observed a significant decrease in *LITATSI* expression in lung adenocarcinoma tissues and mesenchymal breast cancer cells. Moreover, *LITATSI* is localized in the cytoplasm, and its expression can be induced by TGF- β . The remaining question is how *LITATSI* expression is modulated at other levels during cancer progression. Tumor-inhibitory *miR-22-3p* was identified as an upstream modulator of *LITATSI* expression in oral squamous cell carcinoma cells⁶⁷. Thus, specific tumor-promoting miRNAs may target *LITATSI* for degradation in breast cancer and lung cancer progression. Moreover, *LITATSI* was shown to be a short-lived lncRNA that is degraded by nuclear RNases in HepG2 cells⁶⁸. Therefore, certain cancer-related cytoplasmic RNases or RNA-binding proteins may alter *LITATSI* stability. Additionally, we could not rule out the possibility that *LITATSI* expression is changed by epigenetic modifications such as promoter hypermethylation.

Our results showed that higher *LITATSI* expression correlates with a favorable survival outcome in breast and non-small cell lung cancer patients. These results highlight the predictive potential of *LITATSI* expression for cancer progression. Given their cell/tissue-specific expression pattern and dysregulation during cancer progression, lncRNAs are emerging as effective biomarkers for cancers⁶⁹. For example, the urine-based test for the lncRNA *PCA3* has been approved by the FDA for prostate cancer diagnosis⁷⁰ and has been further developed as a promising prognostic marker^{71, 72}. We also found that reintroducing *LITATSI* into highly aggressive MDA-MB-231 cells impaired their migration and extravasation, indicating that *LITATSI* may be a therapeutic agent for cancers. Hence, RNA delivery systems, such as lipid nanoparticles (LNPs), which have been extensively tested and optimized as carriers of therapeutic mRNA molecules^{73, 74}, can be applied to transduce *LITATSI* and evaluate its therapeutic value in preclinical *in vivo* models.

Materials and Methods

Cell culture and reagents

HEK293T (CRL-1573), HepG2 (HB-8065), A549 (CRM-CCL-185), A549 (CCL-185EMT), MDA-MB-231 (CRM-HTB-26), MDA-MB-436 (HTB-130), and MCF7 (HTB-22) cells were purchased from the American Type Culture Collection (ATCC) and cultured in Dulbecco's modified Eagle medium (DMEM; Thermo Fisher Scientific; 41965062) supplemented with 10% fetal bovine serum (FBS; Thermo Fisher Scientific; 16000044) and 100 U/ml penicillin/streptomycin (Thermo Fisher Scientific; 15140163). MCF10A-M1 and MCF10A-M2 cells were kindly provided by Dr. Fred Miller (Barbara Ann Karmanos Cancer Institute, Detroit, USA) and cultured in DMEM/F12 (GlutaMAX™ Supplement; Thermo Fisher Scientific; 31331028) containing 5% horse serum (Thermo Fisher Scientific; 26050088), 0.1 μ g/ml cholera toxin (Sigma–Aldrich; C8052), 0.02 μ g/ml Epidermal Growth Factor (EGF; Sigma–Aldrich; 01-107), 0.5 μ g/ml hydrocortisone (Sigma–Aldrich; H0135), 10 μ g/ml insulin (Sigma–Aldrich; I6634), and 100 U/ml penicillin/streptomycin. All cell lines were maintained in a 5% CO₂, 37 °C humidified incubator, tested monthly for mycoplasma contamination, and checked for authenticity by short tandem repeat (STR) profiling. The protein synthesis inhibitor cycloheximide (CHX; Sigma–Aldrich; C1988) was added to the medium at a concentration of 50 μ g/ml. Two lysosome inhibitors, BafA1 (Sigma–Aldrich; B1793) and HCQ (Sigma–Aldrich; H0915), were used at final concentrations of 20 nM and 20 μ M, respectively. The proteasome inhibitor MG132 (Sigma–Aldrich; 474787) was used at a final concentration of 5 μ M. A selective small molecule kinase inhibitor of T β RI (SB431542; SB)⁷⁵ was used at a concentration of 10 μ M. Recombinant TGF- β 3 and recombinant BMP6 were kind gifts from Andrew Hinck (University of Pittsburgh) and Slobodan Vukicevic (University of Zagreb, Croatia), respectively.

Plasmid construction

Full-length *LITATSI* was amplified by PCR from MDA-MB-231 cell-derived cDNA and inserted into the lentiviral vector pCDH-EF1 α -MCS-polyA-PURO. Two independent shRNAs, CRISPRa gRNAs and CasRx gRNAs were designed and inserted into the lentiviral vectors pLKO.1, lenti sgRNA (MS2)_puro optimized backbone (Addgene; 73797), and pRX004-pregRNA (Addgene; 109054), respectively. *LITATSI* promoter fragments were amplified from MDA-MB-231 genomic DNA and subcloned into the pGL4-luc backbone (Promega). The construct expressing dCasRx-TurboID was modified from a CARPID dCasRx-BASU plasmid (Addgene; 153303) by replacing a fragment expressing *Bacillus subtilis* biotin ligase (BASU) with a fragment expressing TurboID. All plasmids were verified by Sanger sequencing, and the primers used for plasmid construction are listed in Appendix Table S2.

Lentiviral transduction and transfection

To produce lentivirus, packaging plasmids (VSV, gag, and Rev) and expression constructs for cDNAs or shRNAs were cotransfected into HEK293T cells. At 48 h post-transfection, supernatants were collected from HEK293T cells and added to target cells supplemented with the same volume of fresh medium. After 48 h of infection, puromycin (1 μ g/ml; Sigma-Aldrich; P9620) was added to the medium to select stable cells. We used TRCN0000040031 for *SMAD4* knockdown, TRCN0000003478 (#1) and TRCN0000010792 (#2) for *SMURF2* knockdown, TRCN0000127698 (#1) and TRCN0000128209 (#2) for *ZC3H12A* knockdown, and TRCN0000039773 for *TBRI* knockdown. For the transfection of GapmeRs (Eurogentec), 1.2×10^5 A549 cells were seeded in wells of a 12-well plate and incubated with the complex formed by Lipofectamine 3000 (Thermo Fisher Scientific; L3000015) and GapmeRs (25 nM at final concentration). Medium was changed after 6 h. RNA and protein samples were collected at 24 h post-transfection. The sequences of GapmeRs are listed in Appendix Table S3. For siRNA transfection, 10 nM nontargeting siRNA (Dharmacon) or SMARTpool siRNA targeting *SMAD7* (Dharmacon; L-020068-00-0005) was transfected into MDA-MB-231 cells at 80% confluence with DharmaFECT transfection reagents. The medium was changed at 24 h post-transfection.

RT-qPCR

A NucleoSpin RNA kit (Macherey Nagel; 740955) was used to isolate total RNA from cells. Reverse transcription was carried out with a RevertAid RT Reverse Transcription Kit (Thermo Fisher Scientific; K1691). The indicated genes were amplified using the synthesized cDNA with specific primer pairs, and signals were visualized with a CFX Connect Real-Time PCR Detection System (Bio-Rad). *GAPDH* was used as the reference gene for normalization by the $2^{-\Delta\Delta C_t}$ method. The primer sequences used for RT-qPCR are listed in Appendix Table S4.

Western blotting

RIPA buffer (150 mM sodium chloride, 1.0% Triton-X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS), and 50 mM Tris-HCl (pH 8.0)) supplemented with complete protease inhibitor cocktail (Roche; 11836153001) was applied to lyse cells. Subsequently, protein concentrations were evaluated with a DCTM protein assay kit (Bio-Rad; 5000111). Next, SDS-polyacrylamide gel electrophoresis (PAGE) was performed, and proteins were then transferred onto a 0.45- μ m polyvinylidene difluoride (PVDF) membrane (Merck Millipore; IPVH00010). Subsequently, the membrane was blocked with 5% nonfat dry milk in Tris-buffered saline (TBS) with 0.1% Tween 20 detergent (TBST) for 1 h at room temperature. After probing the membranes with the corresponding primary and secondary antibodies, images were acquired with a ChemiDoc Imaging System (Bio-Rad). The primary antibodies used for western blotting are listed in Appendix Table S5. Horseradish peroxidase (HRP)-linked anti-

mouse IgG (Sigma–Aldrich; NA931V) and anti-rabbit IgG (Cell Signaling; 7074S) were used as secondary antibodies. ImageJ (National Institutes of Health, United States) was used to quantify relative protein expression levels by densitometry.

Transcriptional reporter assays

To quantify SMAD3/4-driven transcriptional CAGA-luc reporter activity, 3×10^5 HepG2 cells were seeded in the wells of a 24-well plate. The next day, 100 ng of the SMAD3/4-driven transcriptional CAGA-luc plasmid⁴⁰, 80 ng of the β -galactosidase expression construct, and 320 ng of the indicated expression constructs were cotransfected into HepG2 cells using polyethyleneimine (PEI; Polysciences; 23966). After 16 h incubation and serum starvation for 6–8 h, the cells were stimulated with or without TGF- β (1 ng/ml) for 16 h. To measure the activity of the *LITATSI* promoter fragments, 250 ng of the *LITATSI* promoter luciferase reporter was cotransfected with 80 ng of the β -galactosidase expression construct into HepG2 cells in the presence of PEI or into A549 cells by Lipofectamine 3000. After 16 h incubation and serum starvation for 6–8 h, the cells were stimulated with ligand buffer (vehicle control), TGF- β (5 ng/ml), or BMP6 (50 ng/ml) for 16 h. Luciferase activity was measured with the substrate D-luciferin (Promega) and a luminometer (PerkinElmer) and normalized to β -galactosidase activity. All experiments were performed three times, and representative results are shown.

Immunofluorescence staining

To evaluate the expression and localization of SMURF2 (endogenous or MYC-tagged), immunofluorescence staining was performed as previously described⁷⁶. In brief, cells were fixed with 4% paraformaldehyde (PFA) for 20 min and permeabilized with 0.1% Triton-X in PBS for 10 min. Subsequently, 3% bovine serum albumin (BSA) in PBS was added to block nonspecific binding. For detection of SMURF2, cells were incubated first with a primary antibody against SMURF2 (1:100 dilution; Santa Cruz; sc-393848) or MYC (1:100 dilution; Santa Cruz; sc40) for 45 min at room temperature and then with a secondary antibody (Invitrogen; A21428) for 1 h at room temperature. For F-actin immunofluorescent staining, cells were incubated with Phalloidin conjugated with Alexa Fluor 488 (1:500 dilution; Thermo Fisher Scientific; A12379) for 30 min at room temperature as described before⁷⁷. VECTASHIELD Antifade Mounting Medium with DAPI (Vector Laboratories; H-1200) was used to mount coverslips, and images were acquired with a Leica SP8 confocal microscope (Leica Microsystems).

Ubiquitination assay

HEK293T cells transfected with the indicated constructs and stable MDA-MB-231-HA-Ub cells were treated with 5 μ M MG132 for 5 h prior to harvesting. Cells were lysed in 1% SDS–RIPA buffer (25 mM Tris–HCl (pH 7.4), 150 mM NaCl, 1% NP40, 0.5% sodium deoxycholate, and 1% SDS) supplemented with a protease inhibitor and 10 mM N-ethylmaleimide (NEM; Sigma–Aldrich; E3876). After the lysates were boiled for 5 min and diluted to an SDS concentration of 0.1%, 20 μ l of anti-FLAG agarose (Sigma–Aldrich; A2220) was added to the lysates containing equal amounts of protein and incubated for 30 min at 4°C. To detect the polyubiquitination of endogenous T β RI, cell lysates were incubated with 5 μ l of an antibody against T β RI (Santa Cruz; sc-398) for 16 h at 4°C. The mixture was then incubated with 20 μ l of Protein A Sepharose (GE Healthcare; 17-0963-03) for 2 h at 4°C. After five washes, the beads were boiled in 2 \times sample buffer and analyzed by western blotting.

IncuCyte migration assays

For the wound healing migration assay, 5×10^4 MDA-MB-231 and A549 cells were seeded in the wells of an Essen ImageLock plate (Essen BioScience; 4379). After 16 h culture, the

medium was replaced with DMEM supplemented with 0.5% FBS for another 8 h of culture. A WoundMaker tool (Essen BioScience) was used to generate scratch wounds, after which floating cells were washed away with PBS. An IncuCyte live cell imaging system (Essen BioScience) was used to monitor cell migration. For the chemotactic migration assay, 1×10^3 MDA-MB-231 or A549 cells in DMEM supplemented with 0.5% FBS were seeded in the upper chambers of an IncuCyte Clearview 96-well plate (Essen BioScience; 4582). Then, 200 μ l of DMEM supplemented with 10% FBS was added to the lower reservoir plate. Cells in the top and bottom chambers were imaged and quantified with the IncuCyte system.

Subcellular fractionation

Cells from a 10 cm dish were collected and lysed in 250 μ l of buffer A (50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1% NP40, and 0.25% sodium deoxycholate) for 15 min on ice. After centrifugation at 3,000 g for 5 min, the supernatant was collected and saved as the cytoplasmic fraction. The pellet was washed with PBS twice and resuspended in 150 μ l of buffer B (50 mM Tris-HCl (pH 7.4), 400 mM NaCl, 1% NP40, 0.5% sodium deoxycholate, and 1% SDS). After 20 min of incubation on ice and centrifugation at 12,000 g for 15 min, the supernatant was collected and saved as the nuclear fraction. The isolated cytoplasmic and nuclear fractions were used to quantify the expression of lncRNAs by RT-qPCR.

RACE

RACE was performed on A549 cells according to the manufacturer's instructions of a SMARTer RACE 5' / 3' Kit (TaKaRa; 634859). In brief, 5' and 3' RACE were carried out with specific primers on synthesized cDNA from A549 cells. After agarose gel electrophoresis, DNA was isolated and subcloned into the pRACE vector. Sanger sequencing was performed to analyze the sequence amplified from RACE.

RIP

To identify interactions between lncRNAs and proteins of interest, RIP was performed with a Magna RIPTM RNA-Binding Protein Immunoprecipitation Kit (Merck Millipore; 17-700). In brief, cells were collected and lysed in RIP lysis buffer. After centrifugation at 12,000 g for 10 min at 4°C, the supernatant was collected and supplemented with 700 μ l of wash buffer and 50 μ l of magnetic beads. After being precleared for 6 h at 4°C, the cell lysate was transferred to a new Eppendorf tube with 2.5 μ g of an anti-FLAG antibody (Sigma-Aldrich; F1804), anti-SMURF2 antibody (Santa Cruz, sc-393848), anti-T β RI antibody (Santa Cruz, sc-398), or normal mouse/rabbit IgG and incubated for 16 h at 4°C. For *in vitro* RIP, 9 pmol of *in vitro*-transcribed *LITATSI-S* or *LITATSI-AS* was incubated with 1 pmol T β RI-ICD (CARNA BIOSCIENCES; 09-441-20N) for 16 h at 4°C. The beads were blocked with 5 μ l of yeast tRNA (Invitrogen; AM7119) and 5 μ l of BSA (Invitrogen; AM2618) for 2 h at 4°C and were then added to the cell lysates for another 3 h of incubation at 4°C. Then, the beads were treated with 1.5 μ l of DNase I (Roche; 04716728001) for 10 min at 37°C followed by 1.5 μ l of proteinase K (Merck Millipore; 71049) for 20 min at 56°C. RNA was extracted from the beads, and RT-qPCR was performed as mentioned above.

RNA pull-down assay

RNA pull-down assays were performed to identify *in vitro* interactions between lncRNAs and proteins of interest. In brief, a MEGAscript Kit (Thermo Fisher Scientific; AM1334) was used to synthesize antisense and sense *LITATSI* through *in vitro* transcription. Next, RNA was extracted, and 50 pmol of antisense or sense *LITATSI* was biotinylated with an RNA 3' End Desthiobiotinylation Kit (Thermo Fisher Scientific; 20160). The tertiary structure of each lncRNA was recovered by 10 min of incubation at 70°C followed by gradual cooling to room

temperature. HEK293T cell lysates and recombinant FLAG-SMURF1 protein (Sigma–Aldrich; SRP0227) or recombinant FLAG-SMURF2 protein (Sigma–Aldrich; SRP0228) were incubated with biotinylated lncRNA for 16 h at 4°C. Magnetic beads from a Magnetic RNA–Protein Pull-Down Kit (Thermo Fisher Scientific; 20164) were utilized to capture RNA–protein complexes. Proteins were eluted from the beads and analyzed by western blotting.

CARPID

The CARPID approach was utilized to validate interactions between lncRNAs and proteins of interest at the endogenous level. In brief, biotin (Sigma–Aldrich; B4639) was added to the medium at a final concentration of 200 µM to activate biotinylation for 30 min at 37°C. Cells were collected and lysed in TNE lysis buffer (50 mM Tris–HCl (pH 7.4), 1 mM EDTA, 150 mM NaCl, and 1% NP40) on ice for 10 min. Then, 20 µl of NeutrAvidin™ Agarose (Thermo Fisher Scientific; 29200) was added to the cell lysates with the same amount of protein. The beads were washed with TNE buffer 5 times after incubation for 16 h at 4°C and were then boiled for 5 min in 2× sample buffer. Western blotting was carried out to analyze the enrichment of biotinylated proteins.

PLA

To analyze the endogenous interactions between *LITATSI* and TβRI or SMURF2, a PLA was performed. In brief, A549 cells were seeded on coverslips in the wells of a 24-well plate. After serum starvation for 16 h, the cells were stimulated with or without TGF-β (5 ng/ml) for 2 h. Subsequently, the cells were fixed with 4% PFA for 10 min and permeabilized with PBS supplemented with 0.5% Triton-X for 5 min. The cells were then blocked with Duolink® Blocking Solution for 1 h at 37°C and incubated with primary antibodies against TβRI (Santa Cruz; sc-398) and SMURF2 (Santa Cruz; sc-393848) at a 1:500 dilution for 16 h at 4°C. After three washes with wash buffer A (Sigma–Aldrich; DUO82049), the cells were incubated with secondary antibodies conjugated to the PLUS and MINUS PLA probes (Sigma–Aldrich; DUO92001 and DUO92005) for 1 h at 37°C. Then, ligase (Sigma–Aldrich; DUO92008) was added to the cells and incubated for 30 min prior to incubation with Duolink® Polymerase (Sigma–Aldrich; Cat. Nr.: DUO82028) for 90 min at 37°C. After three washes with wash buffer B (Sigma–Aldrich; Cat. Nr.: DUO82048), the samples were mounted with VECTASHIELD Antifade Mounting Medium with DAPI (Vector Laboratories; H-1200), and images were acquired with a Leica SP8 confocal microscope (Leica Microsystems).

Flow cytometry

Vimentin expression in A549-VIM-RFP cells was quantified by RFP-directed flow cytometry as described elsewhere (Wang et al, 2021). In brief, A549-VIM-RFP stable cells were collected, washed with PBS, and resuspended in PBS containing 5% BSA and 2 mM EDTA (pH 8.0). Subsequently, at least 10,000 cells were acquired with a BD LSR II flow cytometer (BD Biosciences), and the results were analyzed with FlowJo 10.5.0 software.

RNA-seq-based transcriptional profiling, pathway enrichment analysis, and GSEA

To screen for lncRNAs induced by TGF-β, MCF10A-M1, MCF10A-M2, and MDA-MB-231 cells (in biological triplicate) were serum starved for 16 h and stimulated with TGF-β (5 ng/ml) for 0, 2, 8, and 24 h. Then, total RNA was isolated from the cells with TRIzol reagent (Thermo Fisher Scientific; 15596026). Libraries were constructed, and transcriptional analysis was performed on the Illumina HiSeq platform (Beijing Genomics Institute (BGI), Shenzhen). Bioinformatic analysis of differentially expressed transcripts was carried out by BGI. To screen for mRNAs affected by *LITATSI*, we transduced A549 cells with constructs expressing two independent shRNAs (sh*LITATSI* #1 and sh*LITATSI* #2) or a nontargeting shRNA (Co.sh).

After oligo(dT) selection and library preparation, the DNBSeg platform (BGI, Hong Kong) was used to perform RNA-seq. RNA-seq files were processed using the open-source BIODWDL RNAseq pipeline v4.0.0 (<https://zenodo.org/record/3975552#.YiBgxIzMKV4>) developed at Leiden University Medical Center (LUMC). This pipeline performs FASTQ preprocessing (including quality control, quality trimming, and adapter clipping), RNA-seq alignment, read quantification, and optional transcript assembly. FastQC was used for QC checks on raw reads. Adapter clipping was performed using Cutadapt (v2.10) with default settings. RNA-seq read alignment was performed using STAR (v2.7.5a) with the GRCh38 human reference genome. Gene reads were quantified using HTSeq-count (v0.12.4) with the “–stranded = no” setting. The Ensembl version 99 gene annotation was used for quantification. Differential gene expression analysis was performed using R (v3.6.3). First, the gene read count matrix was used to calculate the counts per million mapped reads (CPM) per sample for all annotated genes. Genes with a log₂CPM higher than 1 in at least 25% of all samples were retained for downstream analysis. The numbers of retained genes for each comparison were as follows: Co.sh vs. sh*LITATSI* #1, 12,646 genes; Co.sh vs. sh*LITATSI* #2, 12,692 genes; Co.sh –TGF- β vs. Co.sh +TGF- β , 12,858 genes. For differential gene expression analysis, the dgeAnalysis R-Shiny application (<https://github.com/LUMC/dgeAnalysis/tree/v1.3.1>) was used. EdgeR (v3.28.1) with trimmed mean of M values (TMM) normalization was used to perform differential gene expression analysis. The Benjamini–Hochberg false discovery rate (FDR) was computed to adjust the P-values obtained for each differentially expressed gene. Using a cutoff of 0.05 for the adjusted P-values, up- and downregulated genes were identified. The details of up- and downregulated lncRNAs in response to TGF- β stimulation and the differentially expressed genes upon *LITATSI* depletion were shown in Appendix Tables S1 and S6, respectively. In order to investigate which splice variants of *ZC3H12A-DT* were expressed out of seven splice variants annotated in Ensembl gene annotation version 108, we estimated the raw sequencing reads using StringTie (v1.3.6) that can discriminate the seven splice variants. GSEA was performed with GSEA software⁷⁸. The TGF- β (TGFB_UP.V1_UP) gene response signature³⁹ and EMT (GOBP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION; GO: 0001837) gene signature were used to evaluate the correlations between *LITATSI* and TGF- β /SMAD signaling and EMT, respectively.

Differential gene expression and survival analyses based on patient samples

Differential expression of *LITATSI* was analyzed in samples from patients with breast cancer of different subtypes from TCGA and GTEx datasets using the GEPIA2 database⁷⁹. Patient survival analysis was performed on the Kaplan–Meier Plotter website (<https://kmplot.com/analysis/>)⁸⁰. More details about the databases can be found in Appendix Table S7.

In situ hybridization staining

An RNAScope® Multiplex Fluorescent Kit (Advanced Cell Diagnostics; 323100) and an in situ probe for *LITATSI* (Advanced Cell Diagnostics; 835371-C2) were utilized to evaluate the expression and localization of *LITATSI* in A549 and MDA-MB-231 cells. All fluorescence in situ hybridization procedures were carried out strictly according to the manufacturer's instructions. Images were acquired with a DMi8 inverted fluorescence microscope (Leica). To analyze *LITATSI* expression in patient samples, in situ hybridization was performed on tissue microarrays using a 2.5 HD Detection Kit—BROWN (Advanced Cell Diagnostics; 322300) and the same in situ probe mentioned above. A tissue microarray with lung adenocarcinoma and matched lung tissues was purchased from Biomax (LC1504), and a breast cancer tissue microarray was constructed from the ORIGO cohort (Leiden University Medical Center), which includes 175 breast cancer patients. Patients included in this cohort were diagnosed with

a primary breast tumor and treated in the Leiden University Medical Center (LUMC) between 1997 and 2003³⁶. Informed consent was obtained from all patients. All in situ hybridization procedures were carried out strictly following the manufacturer's instructions for the 2.5 HD Detection Kit—BROWN. Images were acquired with a digital slide scanner (Pannoramic 250 Flash III, 3DHISTECH). The staining index was quantified by the following formula: staining intensity (0, no staining; 1, light brown; 2, brown; 3, dark brown) × proportion of positive cells (0, no positive cells; 1, < 10%; 2, 10–50%; 3, > 50%). The scores were given in a blind manner.

Embryonic zebrafish extravasation assay

The experiments were conducted in a licensed establishment for the breeding and use of experimental animals (LU) and subject to internal regulations and guidelines, stating that advice is taken from the animal welfare body to minimize suffering for all experimental animals housed at the facility. The zebrafish assays described are not considered as an animal experiment under the Experiments on Animals Act (Wod, effective 2014), the applicable legislation in the Netherlands in accordance with the European guidelines (EU directive no. 2010/63/EU) regarding the protection of animals used for scientific purposes. Therefore a license specific for these assays on zebrafish larvae (< 5d) was not required. MDA-MB-231 cells labeled with mCherry were injected into the duct of Cuvier of embryos from transgenic zebrafish (fli; EGFP) as previously described⁸¹. After being maintained in 33°C egg water for 5 days, zebrafish embryos were fixed with 4% formaldehyde. An inverted SP5 STED confocal microscope (Leica) was used to visualize the injected cancer cells and zebrafish embryos. At least 30 embryos per group were analyzed. Two independent experiments were performed, and representative results are shown.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 9. The unpaired Student's t-test was used for most analyses, and $P < 0.05$ was considered statistically significant. All measurements in this study were taken from distinct samples.

Data availability

The RNA-seq data from this publication have been deposited to the GEO database and assigned the identifier GSE203119 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE203119>) and GSE198393 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE198393>).

References

1. Pastushenko, I. & Blanpain, C. EMT transition states during tumor progression and metastasis. *Trends Cell Biol.* **29**, 212-226 (2019).
2. Gui, P. & Bivona, T.G. Evolution of metastasis: new tools and insights. *Trends Cancer* **8**, 98-109 (2022).
3. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* **12**, 31-46 (2022).
4. Sha, Y. *et al.* Intermediate cell states in epithelial-to-mesenchymal transition. *Phys. Biol.* **16**, 021001 (2019).
5. Yang, J. *et al.* Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **21**, 341-352 (2020).
6. Shibue, T. & Weinberg, R.A. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* **14**, 611-629 (2017).
7. van Staalduinen, J., Baker, D., ten Dijke, P. & van Dam, H. Epithelial-mesenchymal-transition-inducing transcription factors: new targets for tackling chemoresistance in cancer? *Oncogene* **37**, 6195-6211 (2018).
8. Dongre, A. & Weinberg, R.A. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* **20**, 69-84 (2019).
9. Colak, S. & ten Dijke, P. Targeting TGF- β signaling in cancer. *Trends Cancer* **3**, 56-71 (2017).
10. Hao, Y., Baker, D. & ten Dijke, P. TGF- β -mediated epithelial-mesenchymal transition and cancer metastasis. *Int. J. Mol. Sci.* **20**, 2767 (2019).

11. Hata, A. & Chen, Y.G. TGF- β Signaling from Receptors to SMADs. *Cold Spring Harb. Perspect. Biol.* **8**, a022061 (2016).
12. Tzavlaki, K. & Moustakas, A. TGF- β Signaling. *Biomolecules* **10**, 487(2020).
13. Yan, X., Xiong, X. & Chen, Y.G. Feedback regulation of TGF- β signaling. *Acta. Biochim. Biophys. Sin. (Shanghai)* **50**, 37-50 (2018).
14. Kavsak, P. *et al.* SMAD7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF- β receptor for degradation. *Mol. Cell* **6**, 1365-1375 (2000).
15. Budi, E.H., Duan, D. & Derynck, R. Transforming growth factor- β receptors and SMADs: regulatory complexity and functional versatility. *Trends Cell Biol.* **27**, 658-672 (2017).
16. Mattick, J.S. & Rinn, J.L. Discovery and annotation of long noncoding RNAs. *Nat. Struct. Mol. Biol.* **22**, 5-7 (2015).
17. Palazzo, A.F. & Koonin, E.V. Functional long Non-coding RNAs evolve from junk transcripts. *Cell* **183**, 1151-1161 (2020).
18. Nandwani, A., Rathore, S. & Datta, M. LncRNAs in cancer: Regulatory and therapeutic implications. *Cancer Lett.* **501**, 162-171 (2021).
19. Statello, L., Guo, C.J., Chen, L.L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96-118 (2021).
20. Lin, C. & Yang, L. Long Noncoding RNA in cancer: wiring signaling circuitry. *Trends Cell Biol.* **28**, 287-301 (2018).
21. Tay, Y., Rinn, J. & Pandolfi, P.P. The multilayered complexity of ceRNA crosstalk and competition. *Nature* **505**, 344-352 (2014).
22. Thomson, D.W. & Dinger, M.E. Endogenous microRNA sponges: evidence and controversy. *Nat. Rev. Genet.* **17**, 272-283 (2016).
23. Watanabe, K. *et al.* Mammary morphogenesis and regeneration require the inhibition of EMT at terminal end buds by *Ovo12* transcriptional repressor. *Dev. Cell* **29**, 59-74 (2014).
24. Chung, V.Y. *et al.* The role of *GRHL2* and epigenetic remodeling in epithelial-mesenchymal plasticity in ovarian cancer cells. *Commun. Biol.* **2**, 272 (2019).
25. Lourenco, A.R. *et al.* C/EBP α is crucial determinant of epithelial maintenance by preventing epithelial-to-mesenchymal transition. *Nat. Commun.* **11**, 785 (2020).
26. Fan, C. *et al.* OVOL1 inhibits breast cancer cell invasion by enhancing the degradation of TGF- β type I receptor. *Signal Transduct. Target. Ther.* **7**, 126 (2022).
27. Jarroux, J. *et al.* HOTAIR lncRNA promotes epithelial-mesenchymal transition by redistributing LSD1 at regulatory chromatin regions. *EMBO Rep.* **22**, e50193 (2021).
28. Ma, Q. *et al.* Inducible lncRNA transgenic mice reveal continual role of HOTAIR in promoting breast cancer metastasis. *Elife* **11**, e79126 (2022).
29. Liu, B. *et al.* A cytoplasmic NF- κ B interacting long noncoding RNA blocks I κ B phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* **27**, 370-381 (2015).
30. Lu, Z. *et al.* Long non-coding RNA NKILA inhibits migration and invasion of non-small cell lung cancer via NF-kappaB/Snail pathway. *J. Exp. Clin. Cancer Res* **36**, 54 (2017).
31. Wu, W. *et al.* LncRNA NKILA suppresses TGF- β -induced epithelial-mesenchymal transition by blocking NF- κ B signaling in breast cancer. *Int. J. Cancer* **143**, 2213-2224 (2018).
32. Wang, L. *et al.* CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res.* **41**, e74 (2013).
33. Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012).
34. Consortium, G.T. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580-585 (2013).
35. Parker, J.S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160-1167 (2009).
36. Out, A.A. *et al.* MUTYH gene variants and breast cancer in a Dutch case-control study. *Breast Cancer Res. Treat* **134**, 219-227 (2012).
37. Györfy, B., Surowiak, P., Budczies, J. & Lánczky, A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *Plos One* **8**, e82241 (2014).
38. Györfy, B. Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. *Comput. Struct. Biotechnol. J.* **19**, 4101-4109 (2021).
39. Padua, D. *et al.* TGF β primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**, 66-77 (2008).
40. Dennler, S. *et al.* Direct binding of SMAD3 and SMAD4 to critical TGF β -inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. *EMBO J.* **17**, 3091-3100 (1998).
41. Yi, W. *et al.* CRISPR-assisted detection of RNA-protein interactions in living cells. *Nat. Methods* **17**,

- 685-688 (2020).
42. Nakao, A. *et al.* Identification of SMAD7, a TGF β -inducible antagonist of TGF- β signalling. *Nature* **389**, 631-635 (1997).
 43. Kang, Y., Chen, C.R. & Massague, J. A self-enabling TGF β response coupled to stress signaling: SMAD engages stress response factor ATF3 for Id1 repression in epithelial cells. *Mol. Cell* **11**, 915-926 (2003).
 44. Massague, J. & Gomis, R.R. The logic of TGF β signaling. *FEBS Lett.* **580**, 2811-2820 (2006).
 45. Katsuno, Y., Lamouille, S. & Derynck, R. TGF- β signaling and epithelial-mesenchymal transition in cancer progression. *Curr. Opin. Oncol.* **25**, 76-84 (2013).
 46. Yuan, J.H. *et al.* A long noncoding RNA activated by TGF- β promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell* **25**, 666-681 (2014).
 47. Richards, E.J. *et al.* Long non-coding RNAs (LncRNA) regulated by transforming growth factor (TGF) β : LncRNA-hit-mediated TGF β -induced epithelial to mesenchymal transition in mammary epithelia. *J. Biol. Chem.* **290**, 6857-6867 (2015).
 48. Wang, P. *et al.* Long noncoding RNA lnc-TSI inhibits renal fibrogenesis by negatively regulating the TGF- β /SMAD3 pathway. *Sci. Transl. Med.* **10**, eaat2039 (2018).
 49. Papoutsoglou, P. *et al.* The TGF β 2-AS1 lncRNA Regulates TGF- β Signaling by Modulating Corepressor Activity. *Cell Rep.* **28**, 3182-3198 e3111 (2019).
 50. Sakai, S. *et al.* Long noncoding RNA ELIT-1 acts as a SMAD3 cofactor to facilitate TGF β /SMAD signaling and promote epithelial-mesenchymal transition. *Cancer Res.* **79**, 2821-2838 (2019).
 51. Papoutsoglou, P. & Moustakas, A. Long non-coding RNAs and TGF- β signaling in cancer. *Cancer Sci.* **111**, 2672-2681 (2020).
 52. Xu, L. *et al.* Long non-coding RNA SMASR inhibits the EMT by negatively regulating TGF- β /SMAD signaling pathway in lung cancer. *Oncogene* **40**, 3578-3592 (2021).
 53. Li, Y. *et al.* Long non-coding RNA SBF2-AS1 promotes hepatocellular carcinoma progression through regulation of miR-140-5p-TGFBR1 pathway. *Biochem. Biophys. Res. Commun.* **503**, 2826-2832 (2018).
 54. Li, Y., Zhao, Z., Sun, D. & Li, Y. Novel long noncoding RNA LINC02323 promotes cell growth and migration of ovarian cancer via TGF- β receptor 1 by miR-1343-3p. *J. Clin. Lab Anal.* **35**, e23651 (2021).
 55. Zhou, B., Guo, W., Sun, C., Zhang, B. & Zheng, F. Linc00462 promotes pancreatic cancer invasiveness through the miR-665/TGFBR1-TGFBR2/SMAD2/3 pathway. *Cell Death. Dis.* **9**, 706 (2018).
 56. Jin, J., Jia, Z.H., Luo, X.H. & Zhai, H.F. Long non-coding RNA HOXA11-AS accelerates the progression of keloid formation via miR-124-3p/TGF β R1 axis. *Cell Cycle* **19**, 218-232 (2020).
 57. Yang, G. & Lin, C. Long noncoding RNA SOX2-OT exacerbates hypoxia-induced cardiomyocytes injury by regulating miR-27a-3p/TGF β R1 Axis. *Cardiovasc Ther.* **2020**, 2016259 (2020).
 58. Cheng, D. *et al.* Long noncoding RNA-SNHG20 promotes silica-induced pulmonary fibrosis by miR-490-3p/TGFBR1 axis. *Toxicology* **451**, 152683 (2021).
 59. Hu, H. *et al.* Long non-coding RNA TCONS_00814106 regulates porcine granulosa cell proliferation and apoptosis by sponging miR-1343. *Mol. Cell Endocrinol.* **520**, 111064 (2021).
 60. Qi, J., Wu, Y., Zhang, H. & Liu, Y. LncRNA NORAD regulates scar hypertrophy via miRNA-26a mediating the regulation of TGF β R1/2. *Adv. Clin. Exp. Med.* **30**, 395-403 (2021).
 61. Zhu, L., Liu, Y., Tang, H. & Wang, P. FOXP3 activated-LINC01232 accelerates the stemness of non-small cell lung carcinoma by activating TGF- β signaling pathway and recruiting IGF2BP2 to stabilize TGFBR1. *Exp. Cell Res.* **413**, 113024 (2022).
 62. Tian, H. *et al.* AKT-induced lncRNA VAL promotes EMT-independent metastasis through diminishing Trim16-dependent Vimentin degradation. *Nat. Commun.* **11**, 5127 (2020).
 63. Yoon, J.H. *et al.* Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat. Commun.* **4**, 2939 (2013).
 64. Hu, W., Qin, L., Li, M., Pu, X. & Guo, Y. A structural dissection of protein-RNA interactions based on different RNA base areas of interfaces. *RSC Adv.* **8**, 10582-10592 (2018).
 65. Sanchez de Groot, N. *et al.* RNA structure drives interaction with proteins. *Nat. Commun.* **10**, 3246 (2019).
 66. Kishore, S. *et al.* A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat. Methods* **8**, 559-564 (2011).
 67. Du, Y. *et al.* Long non-coding RNA LINC01137 contributes to oral squamous cell carcinoma development and is negatively regulated by miR-22-3p. *Cell Oncol. (Dordr)* **44**, 595-609 (2021).
 68. Tani, H., Numajiri, A., Aoki, M., Umemura, T. & Nakazato, T. Short-lived long noncoding RNAs as surrogate indicators for chemical stress in HepG2 cells and their degradation by nuclear RNases. *Sci. Rep.* **9**, 20299 (2019).
 69. Beylerli, O., Gareev, I., Sufianov, A., Ilyasova, T. & Guang, Y. Long noncoding RNAs as promising biomarkers in cancer. *Noncoding RNA Res.* **7**, 66-70 (2022).
 70. Groskopf, J. *et al.* APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis

- of prostate cancer. *Clin. Chem.* **52**, 1089-1095 (2006).
71. Ferro, M. *et al.* Improving the prediction of pathologic outcomes in patients undergoing radical prostatectomy: the value of prostate cancer antigen 3 (PCA3), prostate health index (phi) and sarcosine. *Anticancer Res.* **35**, 1017-1023 (2015).
 72. Cantiello, F. *et al.* PHI and PCA3 improve the prognostic performance of PRIAS and Epstein criteria in predicting insignificant prostate cancer in men eligible for active surveillance. *World J. Urol.* **34**, 485-493 (2016).
 73. Lutz, J. *et al.* Unmodified mRNA in LNPs constitutes a competitive technology for prophylactic vaccines. *NPJ Vaccines* **2**, 29 (2017).
 74. Ickenstein, L.M. & Garidel, P. Lipid-based nanoparticle formulations for small molecules and RNA drugs. *Expert Opin. Drug. Deliv.* **16**, 1205-1226 (2019).
 75. Laping, N.J. *et al.* Inhibition of transforming growth factor (TGF)- β 1-induced extracellular matrix with a novel inhibitor of the TGF- β type I receptor kinase activity: SB-431542. *Mol. Pharmacol.* **62**, 58-64 (2002).
 76. Liu, S. *et al.* Deubiquitinase activity profiling identifies UCHL1 as a candidate oncoprotein that promotes TGF β -induced breast cancer metastasis. *Clin. Cancer Res.* **26**, 1460-1473 (2020).
 77. Sinha, A. *et al.* Visualizing dynamic changes during TGF- β -induced epithelial to mesenchymal transition. *Methods Mol. Biol.* **2488**, 47-65 (2022).
 78. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U S A* **102**, 15545-15550 (2005).
 79. Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, W556-W560 (2019).
 80. Lanczky, A. & Györffy, B. Web-based survival analysis tool tailored for medical research (KMplot): development and implementation. *J. Med. Internet Res.* **23**, e27633 (2021).
 81. Ren, J., Liu, S., Cui, C. & ten Dijke, P. Invasive behavior of human breast cancer cells in embryonic zebrafish. *J Vis Exp* **122**, 55459 (2017).
 82. Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R. & Hofacker, I.L. The Vienna RNA Websuite. *Nucleic Acids Res* **36**, W70-W74 (2008).

Expanded view figures

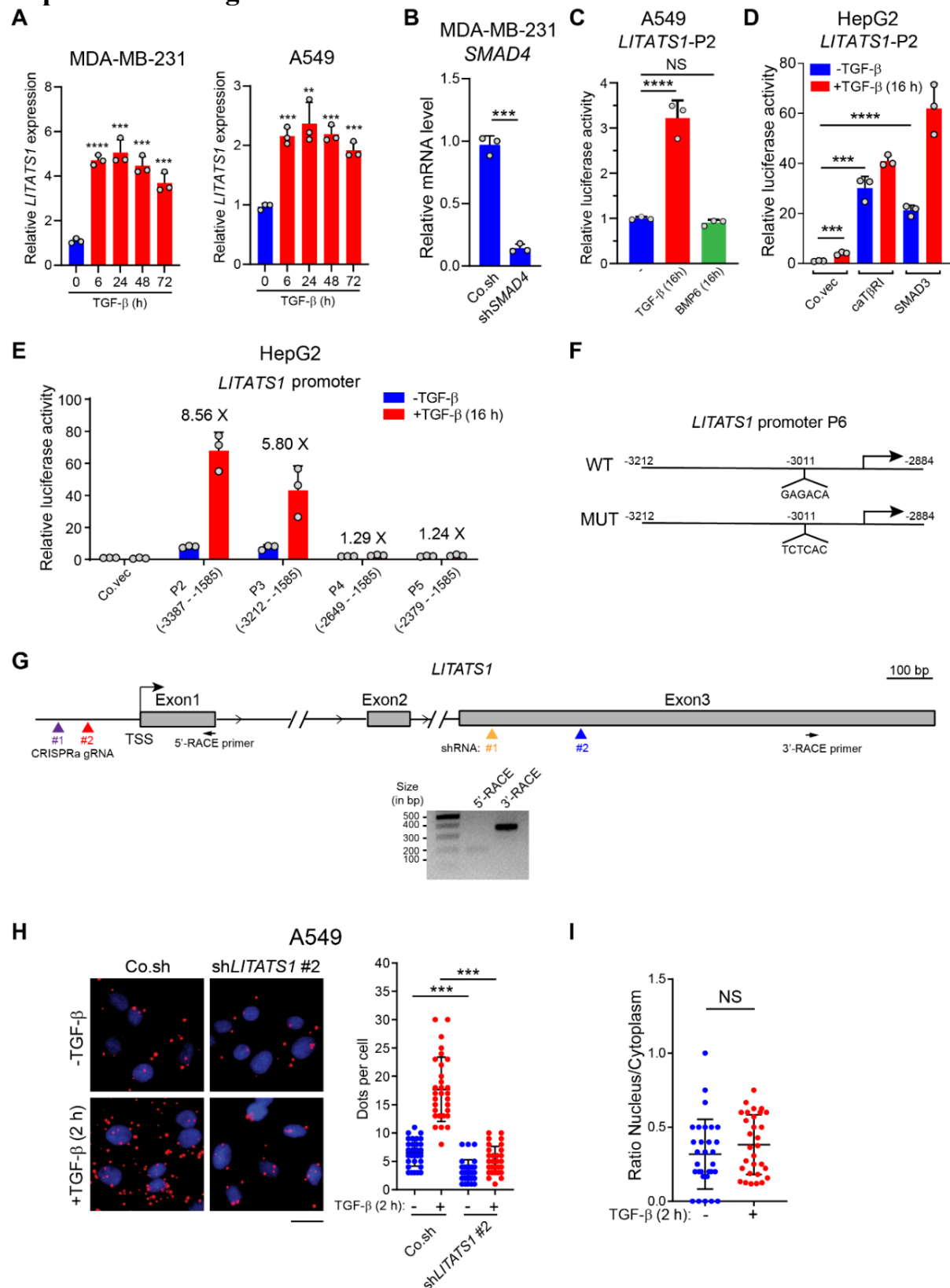


Figure EV1. *LITATS1* is a direct target gene of TGF- β , related to Fig 1. (A) *LITATS1* expression (as analyzed by RT-qPCR) in MDA-MB-231 (left) or A549 (right) cells upon TGF- β stimulation for indicated durations. Representative results from two independent experiments are shown. (B) *SMAD4* expression (as analyzed by RT-qPCR) in MDA-MB-231 cells upon shRNA-mediated *SMAD4* knockdown. Representative results from a minimum of three independent experiments are shown. (C) Analysis of *LITATS1* promoter fragment-mediated transcriptional activity in A549 cells. Cells were transfected with ectopic

LncRNA *LITATS1* suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation

expression constructs for the *LITATS1* promoter 2 luciferase reporter (*LITATS1*-P2) and stimulated with or without TGF- β for 16 h. Representative results from two independent experiments are shown. (D) Effect of caT β RI and SMAD3 on *LITATS1* promoter activity as determined by luciferase reporter assays. HepG2 cells were transfected with ectopic expression constructs for the *LITATS1* promoter 2 luciferase reporter (*LITATS1*-P2) and caT β RI or SMAD3 and were then stimulated with or without TGF- β for 16 h. Representative results from a minimum of three independent experiments are shown. (E) Analysis of *LITATS1* promoter fragment-mediated transcriptional activity in HepG2 cells. Cells were transfected with empty pGL4 vector (Co.vec) or expression constructs for the four indicated *LITATS1* promoter luciferase reporters (P2–P5). Cells were stimulated with or without TGF- β for 16 h. The fold changes between the -TGF- β and +TGF- β groups are indicated. Representative results from a minimum of three independent experiments are shown. (F) Schematic representation of wild-type (WT) and mutant (MUT) *LITATS1* promoter 6 (P6). (G) Schematic illustration of *LITATS1* exons and introns, the target sites of primers for 5' - and 3' -RACE, and the target sites of shRNAs for *LITATS1* knockdown and gRNAs for CRISPRa-mediated *LITATS1* overexpression. TSS: transcription start site. The results of agarose gel electrophoresis of the *LITATS1* 5' - and 3' -RACE DNA products are shown in the lower panel. Representative DNA gel images from two independent experiments are shown. (H) Subcellular distribution of *LITATS1* expression (as detected by RNA fluorescence in situ hybridization) in A549 cells. Cells with or without stable *LITATS1* knockdown were incubated in the presence or absence of TGF- β for 2 h. Representative images are shown in the left panel, and signal quantification data are shown in the right panel. Scale bar = 10 μ m. Representative results from two independent experiments are shown. (I) Lack of effect of TGF- β on the subcellular distribution of *LITATS1* in A549 cells. Quantification of the *LITATS1* nuclear:cytoplasmic signal ratio for each cell was based on the RNA fluorescence in situ hybridization results shown in Fig 1J. Representative results from two independent experiments are shown.

Data information: TGF- β was applied at a final concentration of 5 ng/ml. (A–E) are expressed as the mean \pm SD values from three biological replicates (n=3). (H, I) are expressed as the mean \pm SD values from 30 biological replicates (n=30). **0.001 < P < 0.01; ***0.0001 < P < 0.001; ****P < 0.0001; NS, not significant. Statistical analysis was based on the unpaired Student's t-test.

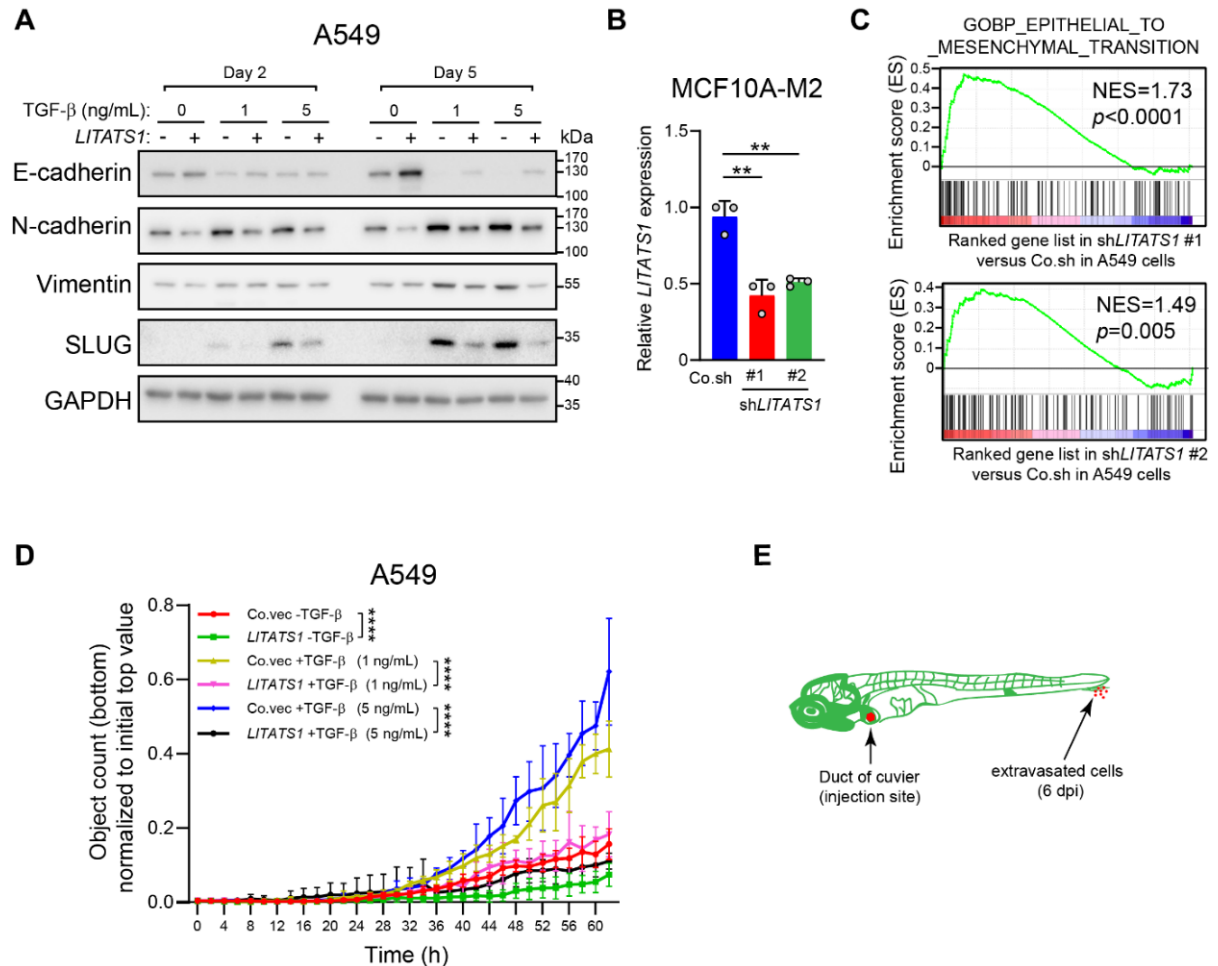


Figure EV2. *LITATS1* inhibits TGF- β -induced EMT and migration in A549 cells, related to Fig 3. (A) Effect of *LITATS1* on TGF- β -induced EMT marker expression in A549 cells upon ectopic *LITATS1* expression. Cells were stimulated without or with different concentrations of TGF- β for 2 or 5 days as indicated. GAPDH, loading control. The results of *LITATS1* overexpression are shown in Appendix Fig S4B. Representative blots from two independent experiments are shown. (B) *LITATS1* expression (as analyzed by RT-qPCR) in MCF10A-M2 cells upon shRNA-mediated knockdown. Representative results from a minimum of three independent experiments are shown. (C) GSEA of positive correlations between (manipulated)

LITATS1 expression and the EMT gene signature. **(D)** An IncuCyte chemotactic migration assay was performed to evaluate the effect of *LITATS1* ectopic expression on TGF- β -induced cell migration in A549 cells. Representative results from two independent experiments are shown. **(E)** Workflow of the breast cancer extravasation experiment in a zebrafish embryo xenograft model. Blood vessels and cancer cells are fluorescently labeled in green and red, respectively.

Data information: **(B, D)** are expressed as the mean \pm SD values from three ($n=3$) and six ($n=6$) biological replicates, respectively. $**0.001 < P < 0.01$; $****P < 0.0001$. In **(B)**, statistical analysis was based on the unpaired Student's t-test. In **(D)**, statistical analysis was based on two-way ANOVA

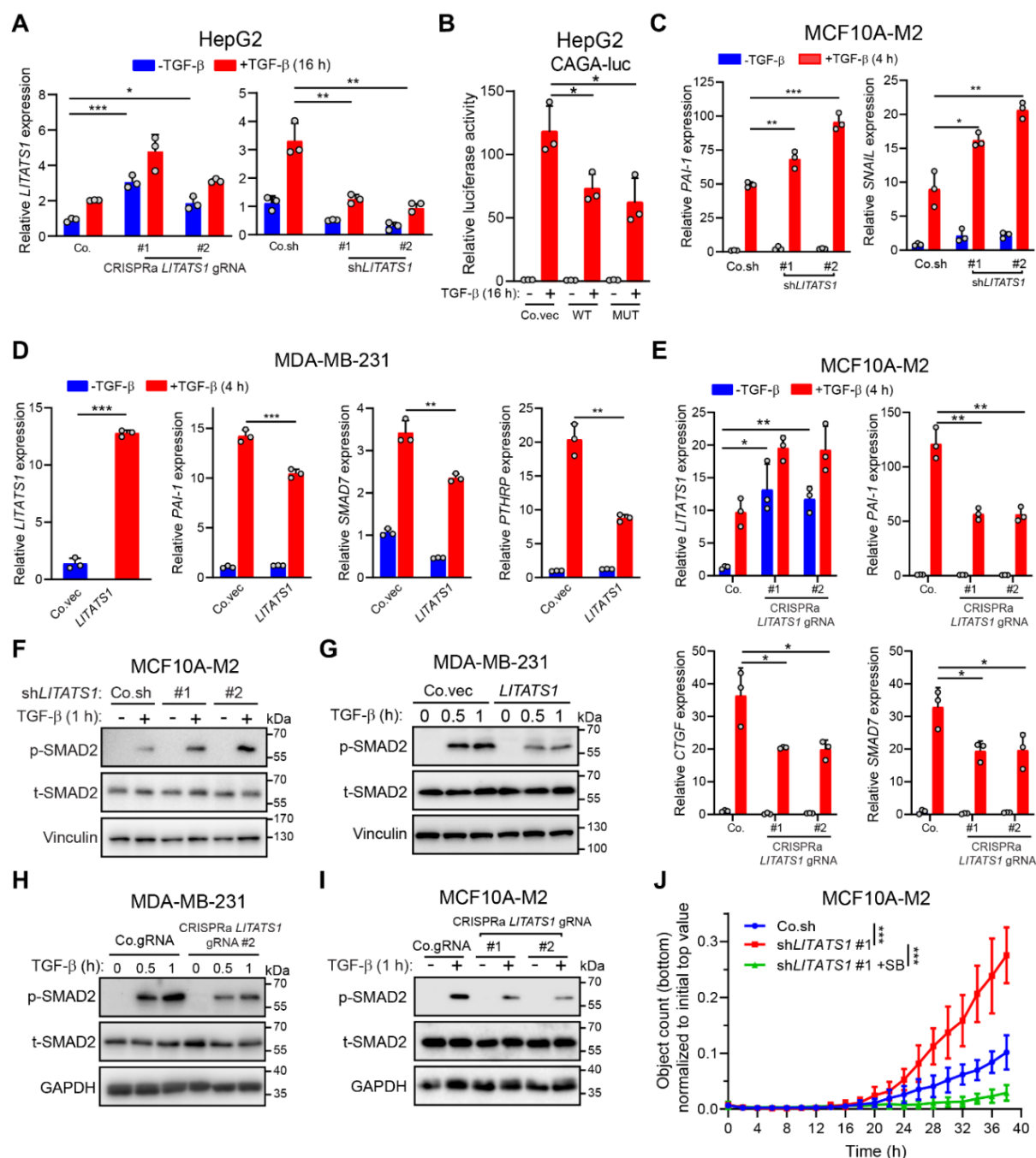


Figure EV3. *LITATS1* suppresses TGF- β /SMAD signaling, related to Fig 4. **(A)** Expression analysis of *LITATS1* (as measured by RT-qPCR) in HepG2 cells upon *LITATS1* overexpression mediated by CRISPRa (left) or *LITATS1* knockdown (right). Cells were transfected with the indicated constructs for *LITATS1* overexpression or depletion and were stimulated with or without TGF- β for 16 h. Representative results from two independent experiments are shown. **(B)** Effect of ectopic expression of wild-type (WT) or mutant (MUT) *LITATS1* on TGF- β /SMAD3 transcriptional activity in HepG2 cells. Cells were transfected with constructs for expression of the TGF- β -induced SMAD3/4-dependent CAGA-luc transcriptional reporter and overexpression of WT or MUT *LITATS1*. Representative results from a minimum of three independent experiments are shown. **(C)** Expression analysis of *PAI-1* and *SNAIL* (as measured by RT-qPCR) in MCF10A-M2 cells upon *LITATS1*

LncRNA *LITATS1* suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation

depletion. Representative results from a minimum of three independent experiments are shown. (D) Expression analysis of *LITATS1*, *PAI-1*, *SMAD7*, and *PTHRP* in MDA-MB-231 cells upon ectopic *LITATS1* expression. Representative results from a minimum of three independent experiments are shown. (E) Expression analysis of *LITATS1*, *PAI-1*, *CTGF*, and *SMAD7* in MCF10A-M2 cells upon CRISPRa-mediated *LITATS1* overexpression. Representative results from a minimum of three independent experiments are shown. (F-I) Effect of *LITATS1* misexpression on the p-SMAD2 level in MCF10A-M2 or MDA-MB-231 cells. The p-SMAD2 and total SMAD2 (t-SMAD2) levels were quantified by western blotting. Vinculin or GAPDH, loading control. Representative blots from a minimum of three independent experiments are shown. (J) An IncuCyte chemotactic migration assay was performed to evaluate the effect of TGF- β signaling inactivation on MCF10A-M2 cell migration mediated by *LITATS1* knockdown. Cells were treated with or without SB431542 (SB; 10 μ M) during the migration assays. Representative results from two independent experiments are shown.

Data information: In (A-I), cells were stimulated with or without TGF- β (1 ng/ml). (A-E) are expressed as the mean \pm SD values from three biological replicates (n = 3). (J) is expressed as the mean \pm SD values from 12 biological replicates (n = 12). *0.01 < P < 0.05; **0.001 < P < 0.01; ***0.0001 < P < 0.001. In (A-E), statistical analysis was based on the unpaired Student's t-test. In (J), statistical analysis was based on two-way ANOVA.

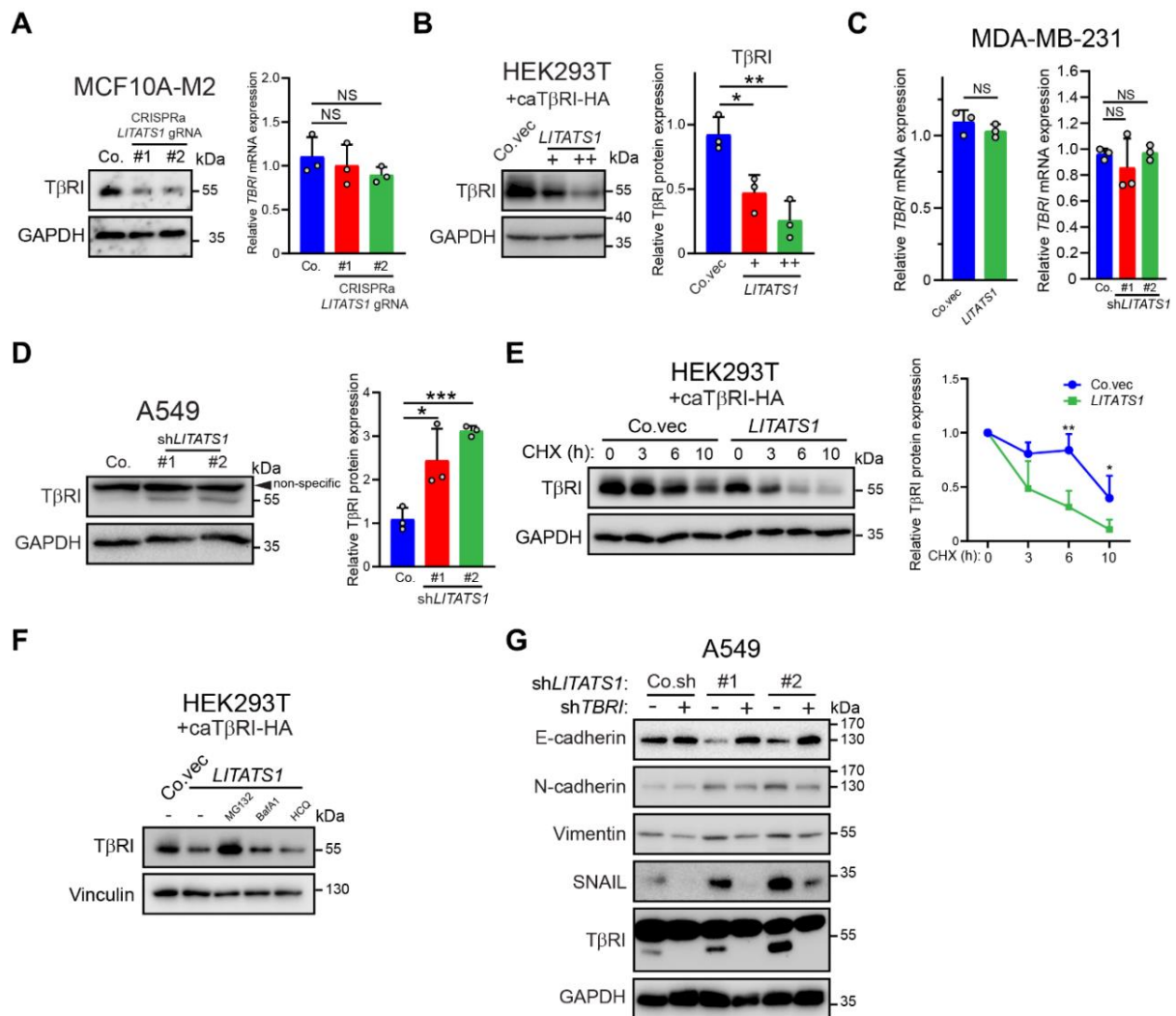


Figure EV4. *LITATS1* enhances T β RI degradation, related to Fig 5. (A) Expression analysis of T β RI in MCF10A-M2 cells upon CRISPRa-induced *LITATS1* overexpression. T β RI protein and T β RI mRNA expressions were evaluated by western blotting (left) and RT-qPCR (right), respectively. GAPDH, loading control. Representative results from a minimum of three independent experiments are shown. (B) Expression analysis of T β RI in HEK293T cells upon ectopic expression of *LITATS1*. Cells were transfected with expression constructs for caT β RI-HA and/or *LITATS1*. Protein expression was evaluated by western blotting (left) and quantification of the relative T β RI protein level is shown in the right panel. GAPDH, loading control. Representative results from a minimum of three independent experiments are shown. (C) Expression analysis of *TβRI* mRNA in MDA-MB-231 cells upon *LITATS1* ectopic expression (left) or *LITATS1* knockdown (right). Representative results from a minimum of three independent experiments are shown. (D) Expression analysis of T β RI protein in A549 cells upon *LITATS1* knockdown. Protein expression was evaluated by western blotting (left) and quantification of the relative T β RI protein level is shown in the right panel. GAPDH, loading control. Representative blots from a minimum of three independent experiments are shown. (E) Effect of CHX on T β RI expression (as measured by western blotting) in HEK293T cells transfected with

Chapter 2

expression constructs for caT β RI-HA and/or *LITATSI*. Cells were treated with CHX (50 μ g/ml) for the indicated durations. GAPDH, loading control. Representative blots from a minimum of three independent experiments are shown. (F) Analysis of T β RI expression (as measured by western blotting) in HEK293T cells transfected with expression constructs for caT β RI-HA and/or *LITATSI* in the presence or absence of the indicated chemical compounds. Cells were incubated with vehicle control (DMSO (-)), a proteasome inhibitor (MG132 (5 μ M)), or a lysosome inhibitor (BafA1 (20 nM) or HCQ (20 μ M)) for 8 h. Vinculin, loading control. Representative blots from a minimum of three independent experiments are shown. (G) Effect of *LITATSI* and/or *TBRI* knockdown on E-cadherin, N-cadherin, Vimentin, SNAIL, and T β RI expression in A549 cells. Cells without (Co.sh) or with stable *LITATSI* knockdown by two independent shRNAs (#1 and #2) were transduced with the *TBRI* shRNA expression construct, and protein expression was analyzed by western blotting. GAPDH, loading control. Representative blots from a minimum of three independent experiments are shown.

Data information: (A–D) are expressed as the mean \pm SD values from three biological replicates (n = 3). (E) is expressed as the mean \pm SD values from four biological replicates (n = 4). *0.01 < P < 0.05; **0.001 < P < 0.01; ***0.0001 < P < 0.001; NS, not significant. Statistical analysis was based on the unpaired Student's t-test.

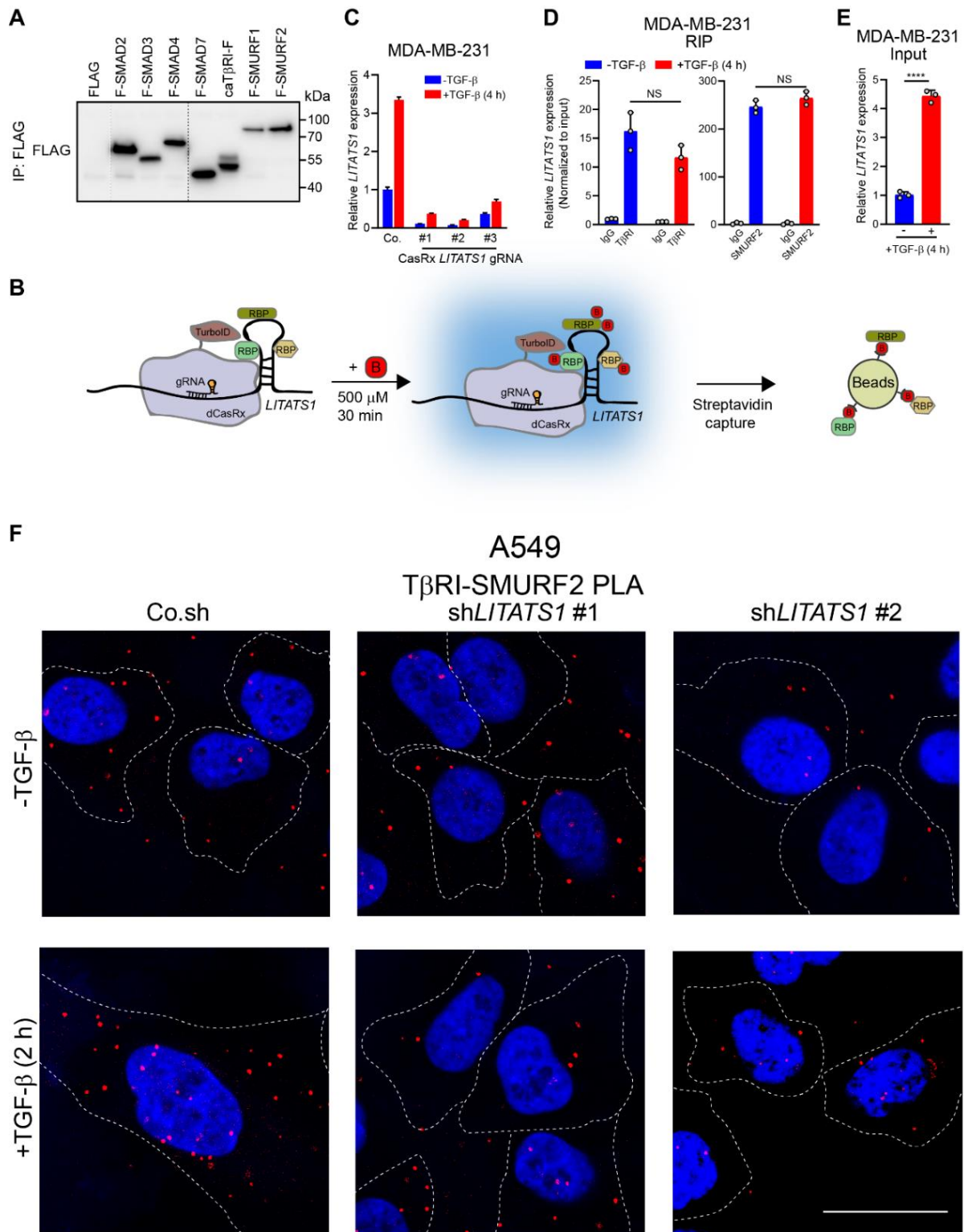
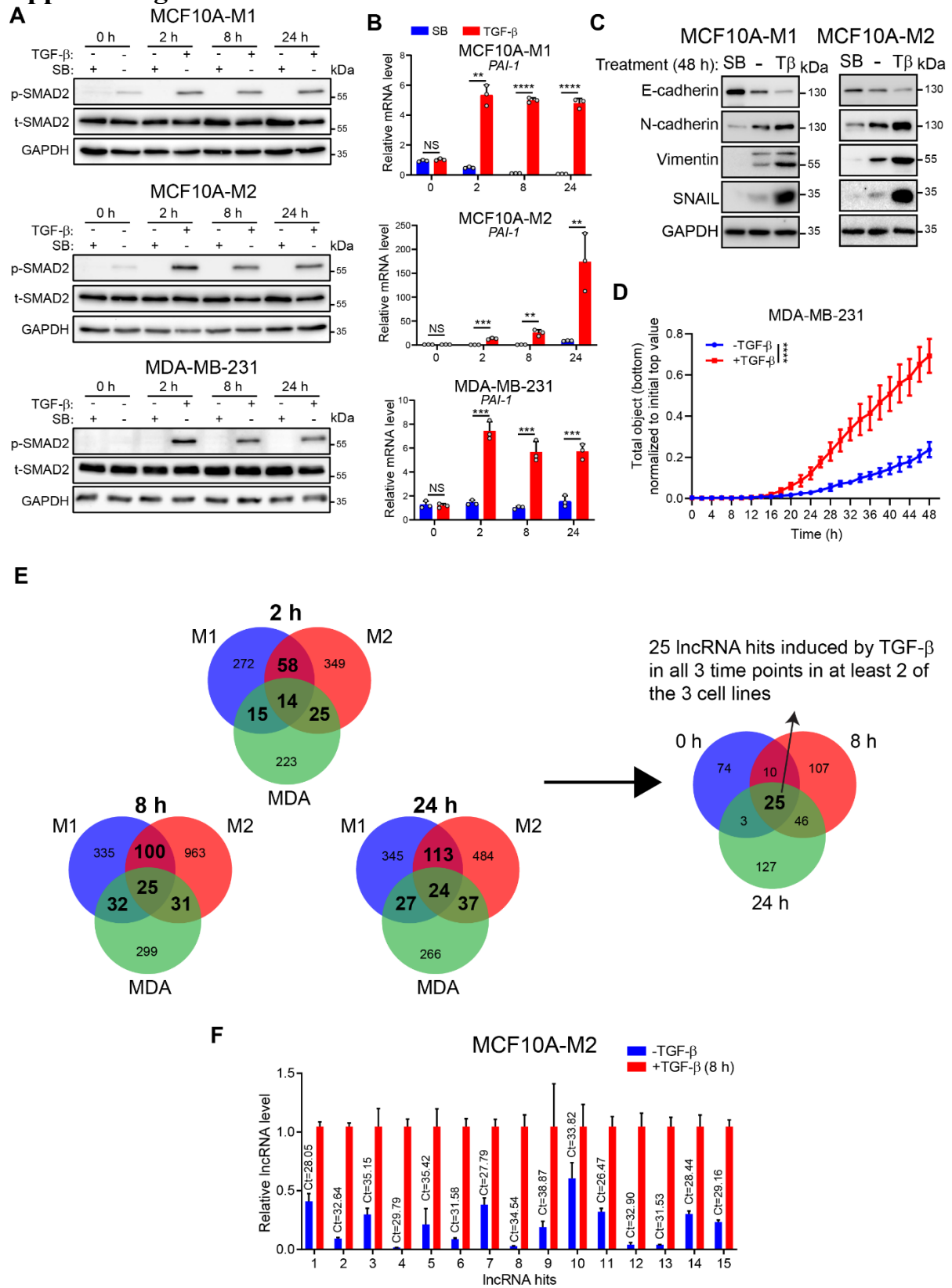


Figure EV5. *LITATS1* interacts with T β RI and SMURF2, related to Fig 6. (A) Expression analysis (by western blotting) of FLAG-tagged proteins in HEK293T cells. Representative blots from two independent experiments are shown. **(B)** Scheme of the CARPID workflow. B: biotin; RBP: RNA-binding protein. **(C)** Expression analysis of *LITATS1* in MDA-MB-231 cells upon CasRx-mediated *LITATS1* knockdown. Stable cells were serum starved for 16 h and stimulated with or without TGF- β (1 ng/ml) for 4 h. The RT-qPCR results are expressed as the mean \pm SEM of technical triplicates. **(D)** Normalization of *LITATS1* expression values in the RIP samples (Fig 6C) to that in the input samples (Fig EV5E). Results are expressed as the mean \pm SD values from three biological replicates ($n = 3$). NS, not significant. Statistical analysis was based on the unpaired Student's t-test. Representative results from two independent experiments are shown. **(E)** Expression analysis of *LITATS1* in the input samples corresponding to Fig 6C. Results are expressed as the mean \pm SD values from three biological replicates ($n = 3$). **** $P < 0.0001$. Statistical analysis was based on the unpaired Student's t-test. Representative results from two

Chapter 2

independent experiments are shown. (F) The endogenous interaction between T β RI and SMURF2 was evaluated by PLA. A549 cells with or without *LITATSI* knockdown were treated with or without TGF- β (5 ng/ml) for 2 h. The red and blue dots indicate the T β RI-SMURF2 interaction and the staining of nuclei by DAPI, respectively. Scale bar = 23.2 μ m. Representative images from two independent experiments are shown.

Appendix figures

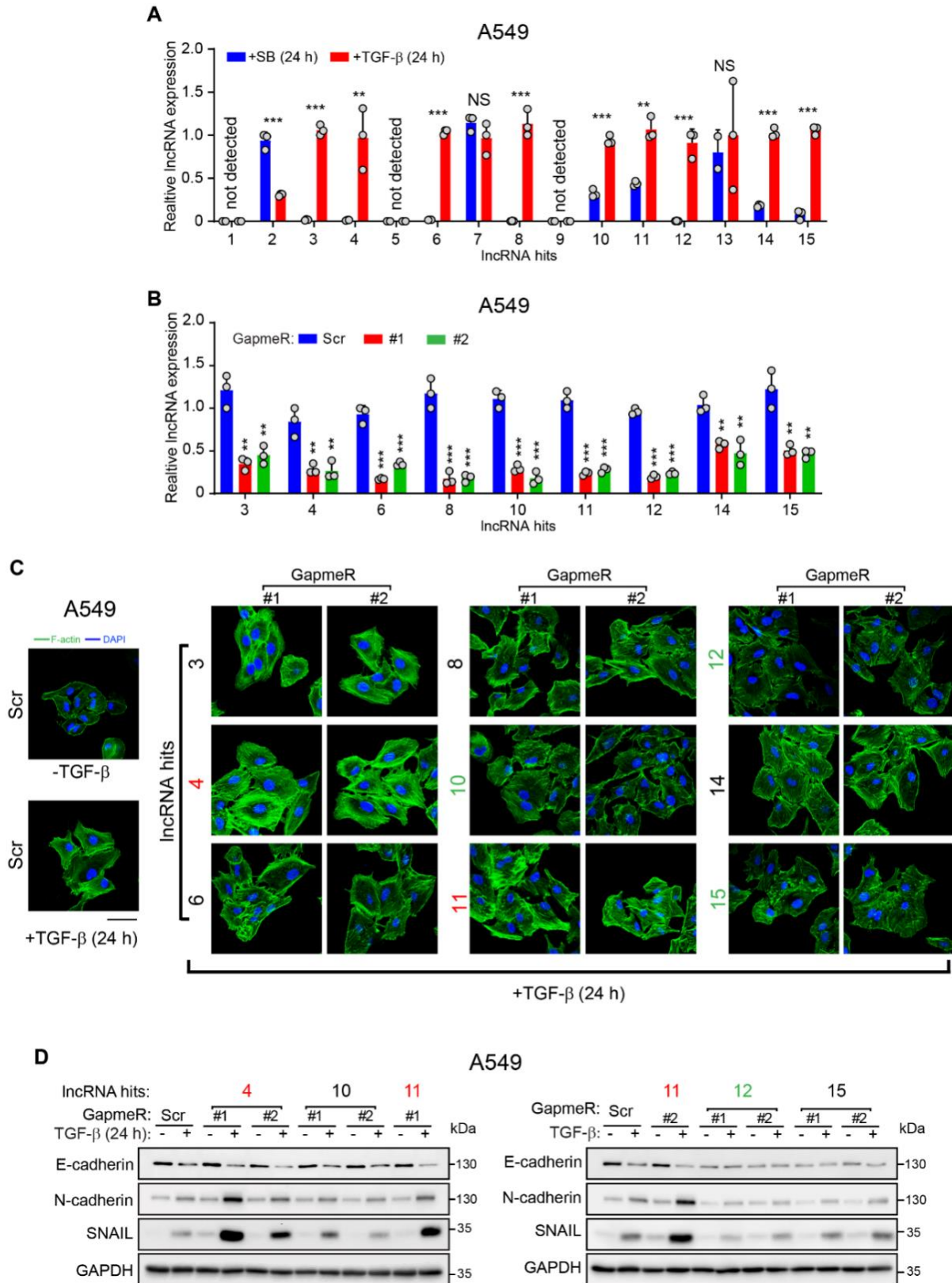


Appendix Fig S1, related to Fig 1. Screening of lncRNAs induced by TGF- β in breast cells. (A) Effect of TGF- β on the levels of p-SMAD2 and t-SMAD2 (as measured by western blotting) in MCF10A-M1 (upper), MCF10A-M2 (middle) and MDA-MB-231 (lower) cells. Cells were treated with SB431542 (SB; 10 μ M) or TGF- β for the indicated durations. GAPDH, loading control. Representative results from a minimum of three independent experiments are shown. (B) Analysis of *PAI-1* mRNA expression in MCF10A-M1 (upper), MCF10A-M2 (middle) and MDA-MB-231 (lower) cells. Cells were treated with

Chapter 2

SB431542 (SB; 10 μ M) or TGF- β for the indicated durations. Representative results from a minimum of three independent experiments are shown. **(C)** Effect of TGF- β stimulation on EMT marker expression in MCF10A-M1 (left) and MCF10A-M2 (right) cells. Cells were stimulated with vehicle control (-), SB431542 (SB; 10 μ M) or TGF- β (Tb) for 48 h, and protein expression was analyzed by western blotting. GAPDH, loading control. Representative results from a minimum of three independent experiments are shown. **(D)** Effect of TGF- β on the migration (as evaluated by an IncuCyte real-time chemotactic migration assay) of MDA-MB-231 cells. Cells were stimulated with or without TGF- β after being seeded in the top chambers. Representative results from a minimum of three independent experiments are shown. **(E)** Twenty-five lncRNAs were significantly induced by TGF- β at all three time points (2 h, 8 h and 24 h) in at least two of the three cell lines (MCF10A-M1 (M1), MCF10A-M2 (M2) and MDA-MB-231 (MDA)). **(F)** Expression analysis of 15 TGF- β -induced lncRNAs in MCF10A-M2 cells. Cells were treated with or without TGF- β for 8 h. The Ct values of the 15 lncRNA hits in cells without TGF- β stimulation are shown. Representative results from a minimum of three independent experiments are shown.

Data information: Cells were stimulated with or without TGF- β (5 ng/mL) for the indicated durations. **(B, D)** are expressed as the mean \pm SD values from three biological replicates ($n = 3$). **(F)** is expressed as the means \pm SD values from three technical replicates ($n = 3$). **, $0.001 < P < 0.01$; ***, $0.0001 < P < 0.001$; ****, $P < 0.0001$; NS, not significant. In **(B)**, statistical analysis was based on unpaired Student's t test. In **(D)**, statistical analysis was based on two-way ANOVA.

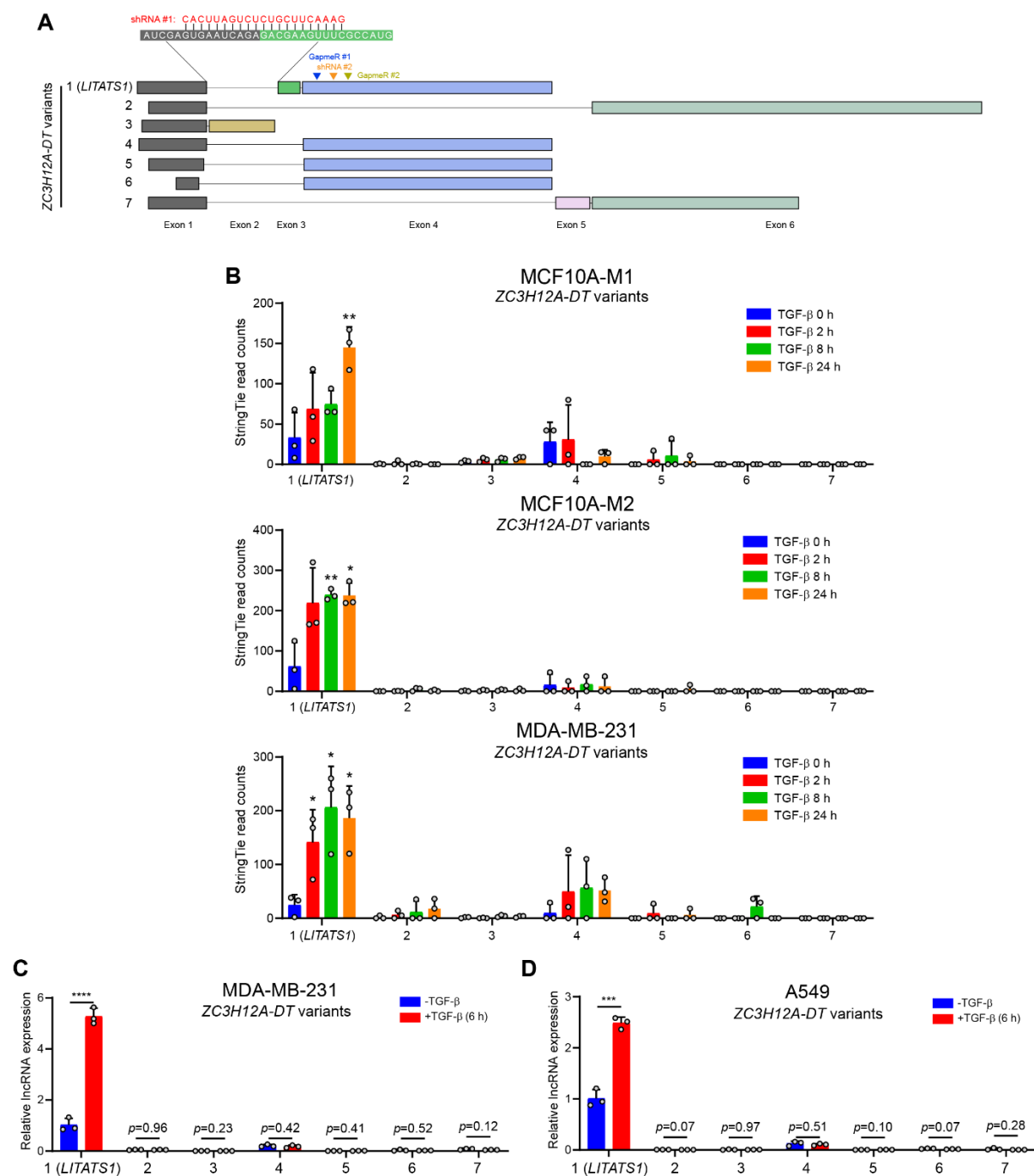


Appendix Fig S2, related to Fig 1. Loss-of-function assays identify *LITATSI* as a suppressor of TGF- β -induced EMT. (A) Expression of 15 lncRNAs (as measured by RT-qPCR) in response to TGF- β in A549 cells. Cells were stimulated with SB431542 (SB; 10 μ M) or TGF- β for 24 h. Representative results from a minimum of three independent experiments are shown. (B) Expression of nine lncRNAs (as measured by RT-qPCR) in A549 cells. Cells were transfected with a scramble GapmeR (Scr) or two independent GapmeRs targeting each hit (i.e., # 1 an #2). Representative results from a minimum of three independent experiments are shown. (C) Effect of lncRNA knockdown on F-actin expression and localization in A549 cells. F-actin was detected by immunofluorescence and nuclei were visualized by DAPI staining. Cells were transfected with a scramble GapmeR (Scr) or two independent GapmeRs targeting each hit (i.e., # 1 an #2) and then stimulated with TGF- β for

Chapter 2

24 h. Scr GapmeR transfected cells without TGF- β stimulation were set as the control. LncRNAs mitigating or potentiating TGF- β -induced F-actin rearrangement are marked in red and green, respectively. Bar=50 μ m. Representative results from a minimum of three independent experiments are shown. **(D)** Effect of lncRNA depletion on EMT marker expression (as detected by western blotting) in A549 cells. Cells were transfected with a scramble GapmeR (Scr) or two GapmeRs targeting each lncRNA hit (i.e., # 1 and #2) and then stimulated without or with TGF- β (1 ng/mL) for 24 h. LncRNAs mitigating or potentiating TGF- β -induced F-actin rearrangement are marked in red and green, respectively. GAPDH, loading control. Representative results from a minimum of three independent experiments are shown.

Data information: In **(A, C)**, cells were stimulated with or without TGF- β (5 ng/mL) for the indicated durations. **(A, B)** are expressed as the mean \pm SD values from three biological replicates (n = 3). **, 0.001 < P < 0.01; ***, 0.0001 < P < 0.001; NS, not significant. In **(A, B)**, statistical analysis was based on unpaired Student's t test.

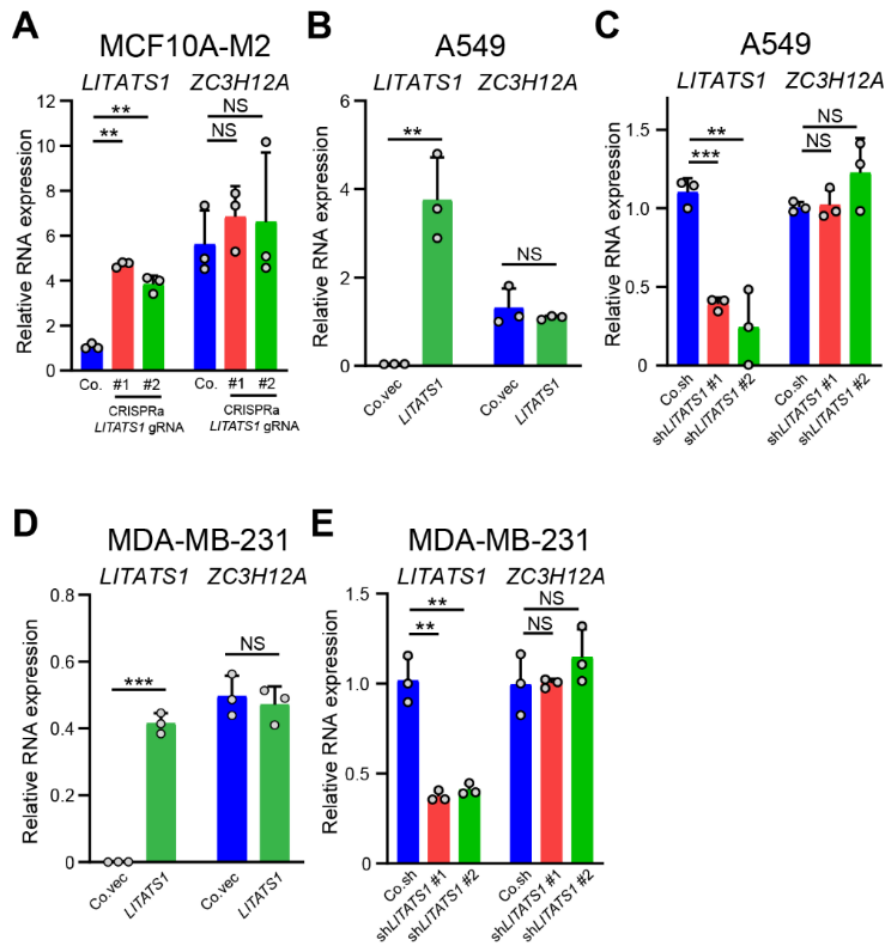


Appendix Fig S3, related to Fig 1. *LITATS1* is the only TGF- β -induced splice variant of *ZC3H12A-DT*. **(A)** Schematic model displays the seven splice variants of *ZC3H12A-DT* and the target sites of shRNAs and GapmeRs for *LITATS1* knockdown. *ZC3H12A-DT* splice variant 1 corresponds to *LITATS1*. **(B)** RNAs-seq read counts of the seven splice variants of *ZC3H12A-DT* including *LITATS1* transcript in breast cell lines. **(C, D)** Expression of the seven *ZC3H12A-DT* splice variants including *LITATS1* transcript in MDA-MB-231 **(C)** and A549 **(D)** cells as measured by RT-qPCR. Cells were stimulated with

LncRNA *LITATS1* suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation

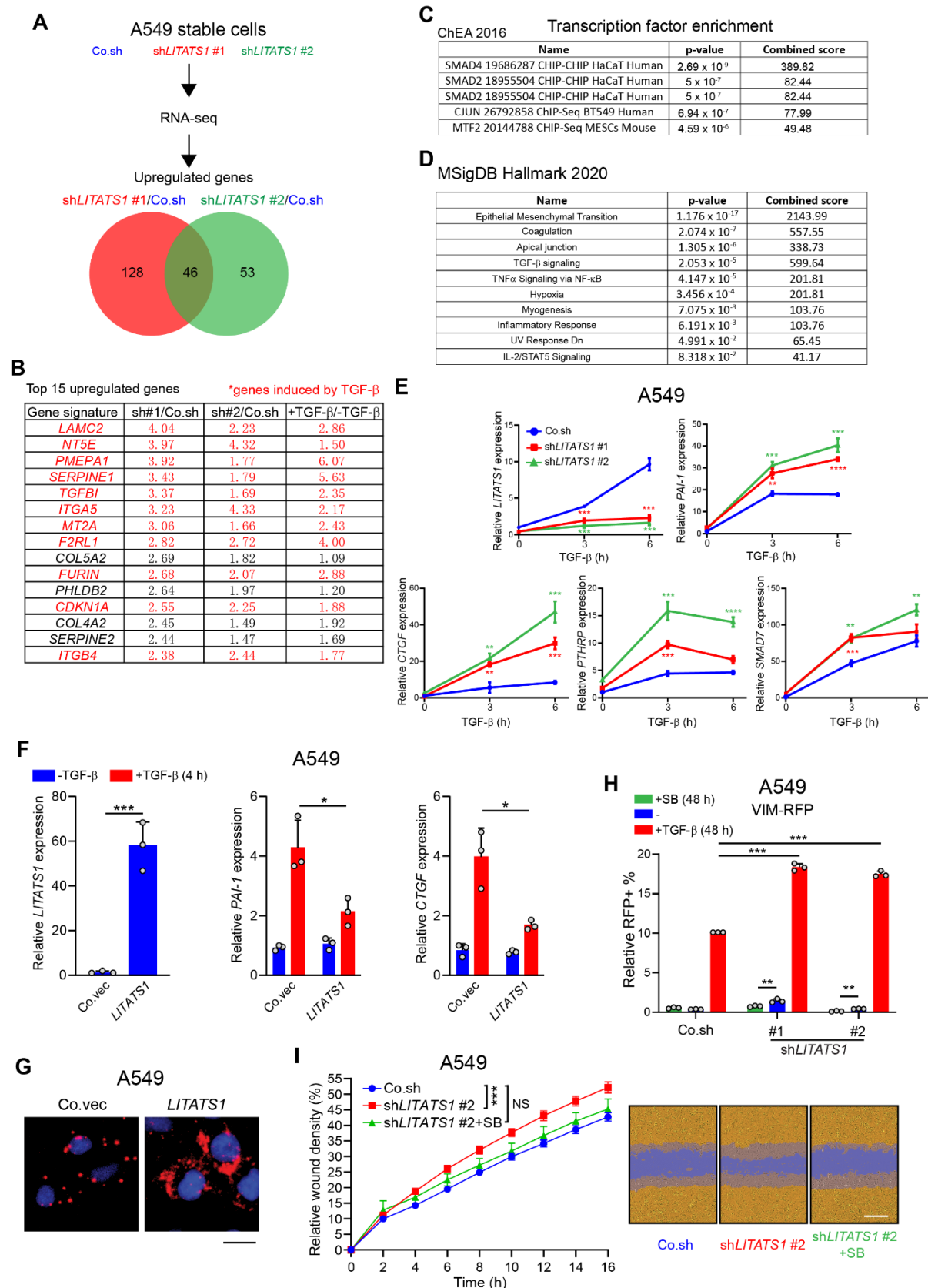
or without TGF- β (5 ng/mL) for 6 h. Representative results from two independent experiments are shown.

Data information: (B-D) are expressed as the mean \pm SD values from three biological replicates (n = 3). *, 0.01 < P < 0.05; **, 0.001 < P < 0.01; ***, 0.0001 < P < 0.001; ****, p < 0.0001. In (B-D), statistical analysis was based on unpaired Student's t test.



Appendix Fig S4, related to Fig 3. *ZC3H12A* expression is not affected by *LITATS1* misexpression. (A-E) Expression of *LITATS1* and *ZC3H12A* mRNA (as measured by RT-qPCR) in MCF10A-M2 (A), A549 (B, C) and MDA-MB-231 (D, E) cells upon *LITATS1* misexpression. Representative results from two independent experiments are shown.

Data information: (A-E) are expressed as the mean \pm SD values from three biological replicates (n = 3). **, 0.001 < P < 0.01; ***, 0.0001 < P < 0.001; NS, not significant. In (A-E), statistical analysis was based on unpaired Student's t test.

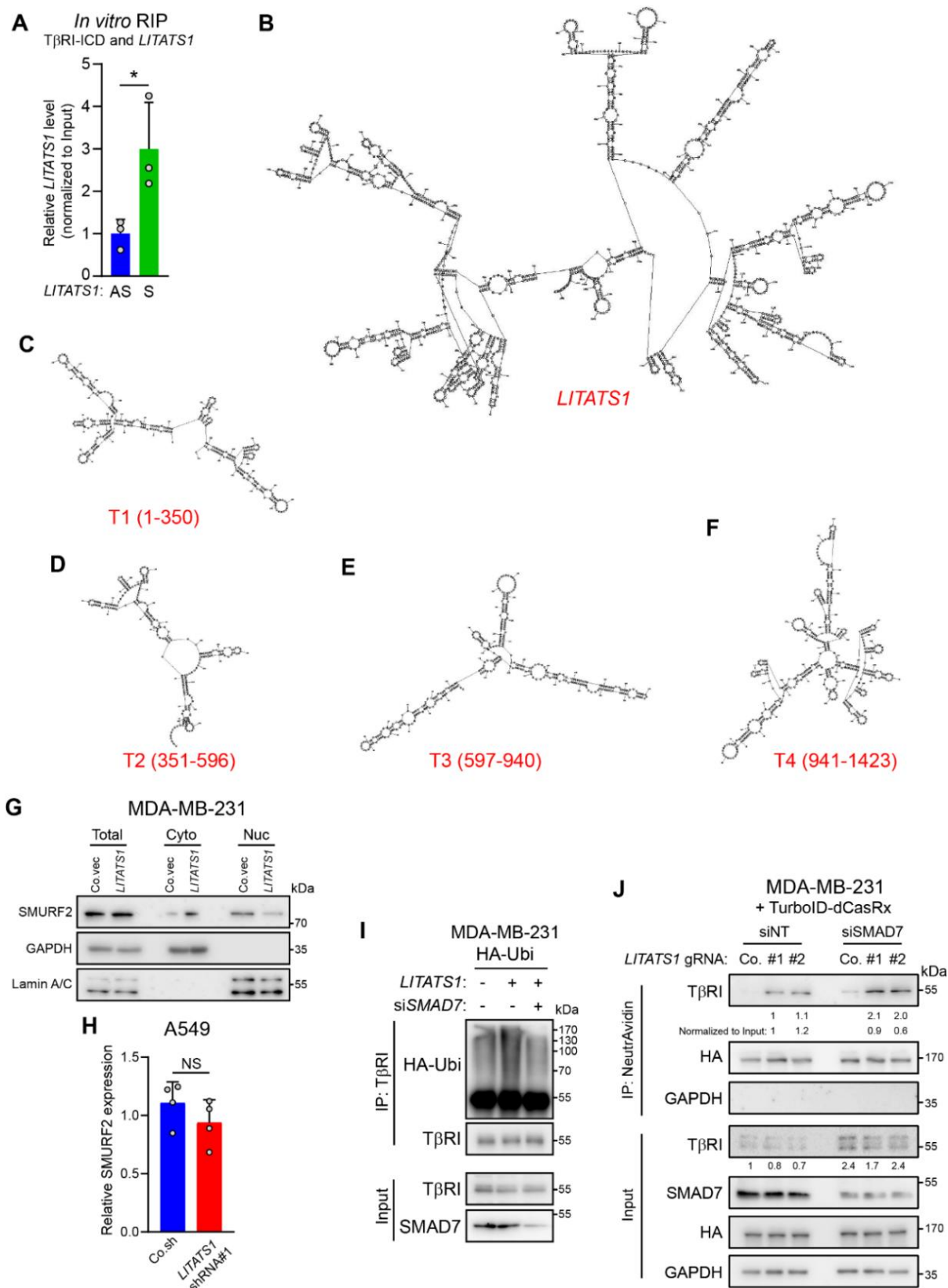


Appendix Fig S5, related to Fig 4. *LITATS1* suppresses TGF- β /SMAD signaling and TGF- β -induced EMT. (A) Scheme for screening genes induced by *LITATS1* knockdown. RNA samples (biological triplicates) were collected from A549 cells with stable *LITATS1* knockdown by two independent shRNAs (#1 and #2). Forty-six upregulated genes (sh*LITATS1* vs. Co.sh) shared by the two shRNAs were identified after analysis of the RNA-seq results. **(B)** The names and fold changes in expression of the top 15 upregulated genes upon *LITATS1* knockdown are listed in the table. RNA-seq results from A549 cells with or without TGF- β (5 ng/mL) stimulation for 2 h as the positive control are shown in the last column. Genes induced by TGF- β

LncRNA *LITATSI* suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation

(fold change > 1.5) are marked in red. **(C, D)** List of transcription factors **(C)** and significantly changed pathways **(D)** affected upon *LITATSI* knockdown. The 46 upregulated genes induced by *LITATSI* knockdown were used as input. **(E)** Expression analysis of *LITATSI*, *PAI-1*, *CTGF*, *PTHRP* and *SMAD7* in A549 cells transduced with shRNAs targeting *LITATSI*. Representative results from a minimum of three independent experiments are shown. **(F)** Expression analysis of *LITATSI*, *PAI-1* and *CTGF* in A549 cells upon *LITATSI* ectopic expression. Representative results from a minimum of three independent experiments are shown. **(G)** Representative images of *LITATSI* expression and localization as detected by RNA fluorescence in situ hybridization in A549 cells. Cells were transduced without (Co.vec) or with *LITATSI* expression constructs. Bar=10 mm. Representative results from two independent experiments are shown. **(H)** Effect of *LITATSI* depletion on TGF- β -induced Vimentin expression in A549 cells. The RFP signal was detected by flow cytometry in A549 VIM-RFP cells in which endogenous VIM was tagged with RFP. Cells with stable *LITATSI* knockdown were treated with vehicle control (-), SB431542 (SB; 10 μ M) or TGF- β (Tb; 5 ng/mL) for 48 h. Representative results from two independent experiments are shown. **(I)** Effect of SB431542 on the migration of A549 cells with or without *LITATSI* knockdown. Cells without (Co.sh) or with stable *LITATSI* knockdown were treated with or without SB431542 (10 μ M) overnight before wounds were generated. The relative wound density (closure) was plotted at the indicated time points (upper). Representative scratch wounds are shown at the experimental endpoint (lower). The regions covered and not covered by migrated cells are colored gray and purple, respectively. Bar=400 μ m. Representative results from two independent experiments are shown.

Data information: In **(E, F)**, cells were serum starved overnight and stimulated with or without TGF- β (1 ng/mL). **(E, F, H)** are expressed as the mean \pm SD values from three biological replicates (n = 3). **(I)** is expressed as the mean \pm SD values from six biological replicates (n = 6). *, 0.01 < P < 0.05; **, 0.001 < P < 0.01; ***, 0.0001 < P < 0.001; ****, p < 0.0001; NS, not significant. In **(E, F, H)**, statistical analysis was based on unpaired Student's t test. In **(I)**, two-way ANOVA was applied to calculate the statistical significance.



Appendix Fig S6, related to Fig 7. *LITATS1* promotes cytoplasmic retention of SMURF2. (A) Interactions between in vitro-transcribed sense (S) or antisense (AS) *LITATS1* and recombinant T β RI-ICD protein were analyzed by RIP. RT-qPCR was performed to detect *LITATS1* expression in immunoprecipitates. The values of *LITATS1* expression were based on normalizing *LITATS1* expression in the RIP samples to that in the input samples and are expressed as the mean \pm SD values from three biological replicates (n = 3). *, 0.01 < P < 0.05. Statistical analysis was based on unpaired Student's t test. Representative results from a minimum of three independent experiments are shown. (B-F) Prediction of the secondary structures of full-length *LITATS1* (B) and truncation mutants (T1-T4; C-F) with the RNAfold software⁸². (G) Effect of *LITATS1*

LncRNA *LITATS1* suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation

overexpression on SMURF2 localization in MDA-MB-231 cells. After subcellular protein fractionation, western blotting was performed to evaluate SMURF2 expression in whole-cell lysates (Total) and the cytoplasmic (Cyto) and nuclear (Nuc) fractions. The levels of the cytoplasmic marker GAPDH and the nuclear marker Lamin A/C are included to demonstrate subcellular protein fractionation. Representative results from a minimum of three independent experiments are shown. **(H)** SMURF2 protein expression is unchanged upon *LITATS1* knockdown in A549 cells. Quantification of western blotting results is expressed as the mean \pm SD values from four biological replicates ($n = 4$). NS, not significant. Statistical analysis was based on unpaired Student's t test. Representative results from a minimum of three independent experiments are shown. **(I)** Effect of SMAD7 knockdown on *LITATS1*-mediated T β RI polyubiquitination. MDA-MB-231 cells with stable HA-Ub expression were transduced with expression constructs for *LITATS1* and/or a *SMAD7* siRNA, as indicated. Polyubiquitination of T β RI was evaluated by western blotting. Representative blots from two independent experiments are shown. **(J)** Interactions between *LITATS1* and T β RI upon *SMAD7* knockdown in MDA-MB-231 cells were detected by the CARPID approach. Cells with stable expression of TurboID-dCasRx were transduced without (Co.) or with (#1 and #2) *LITATS1* targeting gRNAs and the siRNAs targeting *SMAD7*. Cells were stimulated with or without TGF- β (2.5 ng/mL) for 2 h and were then stimulated with biotin (500 mM) for 30 min. Western blotting was performed to detect T β RI expression in whole-cell lysates (Input) and immunoprecipitates (IP). GAPDH and HA-dCasRx expression levels were measured for equal loading of Input samples and as the negative control or positive control, respectively, for proximity biotinylation in immunoprecipitate (IP) samples. Expression of T β RI in the IP samples was normalized to that in the input samples. Representative results from two independent experiments are shown.

Appendix tables are online at <https://www.embopress.org/doi/full/10.15252/embj.2022112806>.



Chapter 3

The lncRNA *LETS1* promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism

Chuannan Fan^{1,2}, Román González-Prieto^{1,3,4}, Thomas B. Kuipers⁵, Alfred C. O. Vertegaal¹, Peter A. van Veelen⁶, Hailiang Mei⁵, and Peter ten Dijke^{1,2*}

¹Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands.

²Oncode Institute, Leiden University Medical Center, Leiden, The Netherlands.

³Genome Proteomics laboratory, Andalusian Center for Molecular Biology and Regenerative Medicine (CABIMER), University of Seville, Seville, Spain.

⁴Department of Cell Biology, University of Seville, Seville, Spain

⁵Department of Biomedical Data Sciences, Sequencing Analysis Support Core, Leiden University Medical Center, Leiden, The Netherlands.

⁶Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands.

*Corresponding author. Tel: +31 71 526 9271; Fax: +31 71 526 8270; E-mail: p.ten_dijke@lumc.nl

Abstract

Transforming growth factor- β (TGF- β) signaling is a critical driver of epithelial-to-mesenchymal transition (EMT) and cancer progression. In SMAD-dependent TGF- β signaling, activation of the TGF- β receptor complex stimulates the phosphorylation of the intracellular receptor-associated SMADs (SMAD2 and SMAD3), which translocate to the nucleus to promote target gene expression. SMAD7 inhibits signaling through the pathway by promoting the polyubiquitination of the TGF- β type I receptor (T β RI). We identified an unannotated nuclear long noncoding RNA (lncRNA) that we designated *LETS1* (lncRNA enforcing TGF- β signaling 1) that was not only increased but also perpetuated by TGF- β signaling. Loss of *LETS1* attenuated TGF- β -induced EMT and migration in breast and lung cancer cells *in vitro* and extravasation of the cells in a zebrafish xenograft model. *LETS1* potentiated TGF- β -SMAD signaling by stabilizing cell surface T β RI, thereby forming a positive feedback loop. Specifically, *LETS1* inhibited T β RI polyubiquitination by binding to nuclear factor of activated T cells (NFAT5) and inducing the expression of the gene encoding the orphan nuclear receptor 4A1 (NR4A1), a component of a destruction complex for SMAD7. Overall, our findings characterize *LETS1* as an EMT-promoting lncRNA that potentiates signaling through TGF- β receptor complexes.

Introduction

Epithelial-to-mesenchymal transition (EMT) is a cellular transdifferentiation process in which epithelial cells lose their cell-cell adhesions and gain the traits of mesenchymal cells¹. This process is characterized by the loss of the epithelial marker E-cadherin and the induction of mesenchymal markers such as N-cadherin and Vimentin. Cancer cells undergoing EMT acquire migratory and invasive properties and become resistant to chemotherapy^{2,3}. Several intermediate states, termed as partial or hybrid EMT states, occur during EMT of cancer cells⁴. Because the process is highly dynamic and reversible, these cancer cells demonstrate a high amount of plasticity and exhibit increased aggressiveness⁴⁻⁷. Moreover, the hybrid EMT RNA signature is correlated with a poor patient prognosis in multiple cancer types⁸⁻¹⁰.

Transforming growth factor- β (TGF- β) signaling plays a crucial role in cancer cell progression through the induction of EMT^{11,12}. Binding of TGF- β ligands enables the TGF- β type II serine-threonine kinase receptor (T β RII) to activate the type I receptor (T β RI), which induces phosphorylation of SMAD2 and SMAD3 (SMAD2/3). Upon forming complexes with SMAD4, activated SMAD2/3 translocate into the nucleus to regulate target gene transcription^{13,14}. TGF- β signaling is tightly controlled at multiple levels^{15,16}. The E3 ligase SMAD ubiquitination regulatory factor 2 (SMURF2) is recruited by inhibitory SMAD7 to target T β RI for polyubiquitination and degradation¹⁷. SMAD7 itself is also fine-tuned by polyubiquitination directed by various E3 ligases, including ARKADIA and ring finger protein 12 (RNF12)^{18,19}. Moreover, the orphan nuclear receptor 4A1 (NR4A1) interacts with complexes composed of AXIN2 and RNF12 or ARKADIA to facilitate SMAD7 polyubiquitination and subsequent proteasomal and lysosomal degradation²⁰.

Long noncoding RNAs (lncRNAs) are emerging as critical players in modulating signaling transduction and cancer progression²¹⁻²³. As a family of noncoding RNAs that are longer than 200 nucleotides in length, lncRNAs can act as scaffolds, guides, or decoys to alter protein-protein interactions or the accessibility of proteins to DNA, thereby enabling them to change signaling transduction at multiple levels^{24,25}. MicroRNAs can be sponged by lncRNAs through the competitive endogenous RNA mechanism^{26,27}. TGF- β -induced responses can be regulated by the induction of certain lncRNAs that serve as effectors^{28,29}. Moreover, the expression or activity of TGF- β signaling components is altered by lncRNAs acting as modulators³⁰⁻³⁴. To

identify additional lncRNAs that participate in TGF- β -induced EMT and cancer progression, we performed a transcriptome screen in three breast cell lines and identified 15 lncRNAs whose expression can be induced by TGF- β -SMAD signaling. One of these TGF- β -induced lncRNAs, *LITATSI*, inhibits TGF- β signaling and TGF- β -induced EMT by promoting the degradation of T β RI³⁵. Here, we focused on an unannotated lncRNA that we named *LETS1* (lncRNA enforcing TGF- β signaling 1), because it promoted TGF- β -SMAD signaling and TGF- β -induced EMT, migration, and extravasation in breast and lung cancer cells. *LETS1* knockdown enhanced polyubiquitination of T β RI. Mechanistically, *LETS1* induced *NR4A1* expression by interacting with nuclear factor of activated T cells (NFAT5) and potentiating NFAT5-mediated *NR4A1* transcription. These findings reveal another layer of T β RI signaling regulation by a previously uncharacterized lncRNA. Targeting *LETS1* may provide a promising therapeutic opportunity to restrain overly active TGF- β signaling in EMT and cancer progression.

Results

LETS1 is a nuclear lncRNA induced by TGF- β -SMAD signaling

We previously reported on lncRNAs that are potently induced by TGF- β by performing transcriptional profiling of three breast cell lines: nonmalignant MCF10A-M1 cells, premalignant MCF10A-M2 cells, and MDA-MB-231 adenocarcinoma cells (fig. S1A)³⁵. In this study, we focused on the TGF- β -induced lncRNA *LETS1* for further investigation (fig. S1A). To characterize *LETS1*, we first confirmed the induction of *LETS1* by TGF- β in A549 lung adenocarcinoma cells and breast cell lines (Fig. 1A). To test whether TGF- β -induced *LETS1* expression was mediated by the canonical SMAD pathway, we knocked down *SMAD2*, *SMAD3*, or *SMAD4* using independent short hairpin RNA(s) [shRNA(s)] in MDA-MB-231 cells (fig. S1B). We observed that TGF- β -induced *LETS1* expression was greatly attenuated upon depletion of *SMAD2*, *SMAD3*, or *SMAD4* (Fig. 1B). Moreover, TGF- β increased *LETS1* expression in MDA-MB-231 cells that were pretreated with cycloheximide (CHX), implying that new protein synthesis was not required for TGF- β to induce *LETS1* expression (fig. S1C). We then mapped the *LETS1* locus on chromosome 15 [chromosome 15: 82098836 to 82101500 (GRCh38.p14)] and revealed that *LETS1* was a single-exon intergenic transcript using 5' and 3' rapid amplification of cDNA ends (RACE) assays (Fig. 1C and fig. S1D). Sequence similarity search by the Basic Local Alignment Search Tool³⁶ showed that the sequence of *LETS1* is unique in the human transcriptome. We evaluated the coding potential of *LETS1* using the Coding Potential Assessment Tool (CPAT)³⁷, which predicted a lack of coding capability for *LETS1* as compared with other protein-coding mRNAs [*ACTB2* and *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase); fig. S1E]. Reverse transcription quantitative polymerase chain reaction (RT-qPCR) after subcellular fractionation in the three breast cell lines revealed that *LETS1* was predominantly localized in the nucleus (Fig. 1D). In addition, fluorescence in situ hybridization using a specific *LETS1* probe showed that TGF- β stimulation enhanced the *LETS1* nuclear signal, which was strongly decreased upon GapmeR-mediated *LETS1* depletion in MDA-MB-231 cells (Fig. 1E). Together, these results demonstrated that *LETS1* is a TGF- β -SMAD-induced lncRNA mainly localized in the nucleus.

LETS1 promotes TGF- β -induced EMT, migration, and extravasation of cancer cells

Because the products of TGF- β -SMAD signaling target genes frequently function as modulators or effectors of TGF- β -SMAD signaling, we determined whether *LETS1* influenced TGF- β -induced EMT in cancer cells. Depletion of *LETS1* transcripts by CRISPR-CasRx attenuated the TGF- β -induced decrease in E-cadherin and increase in N-cadherin, Vimentin, and the EMT-promoting transcription factor SNAIL in A549 cells (Fig. 2A and fig. S2, A and B).

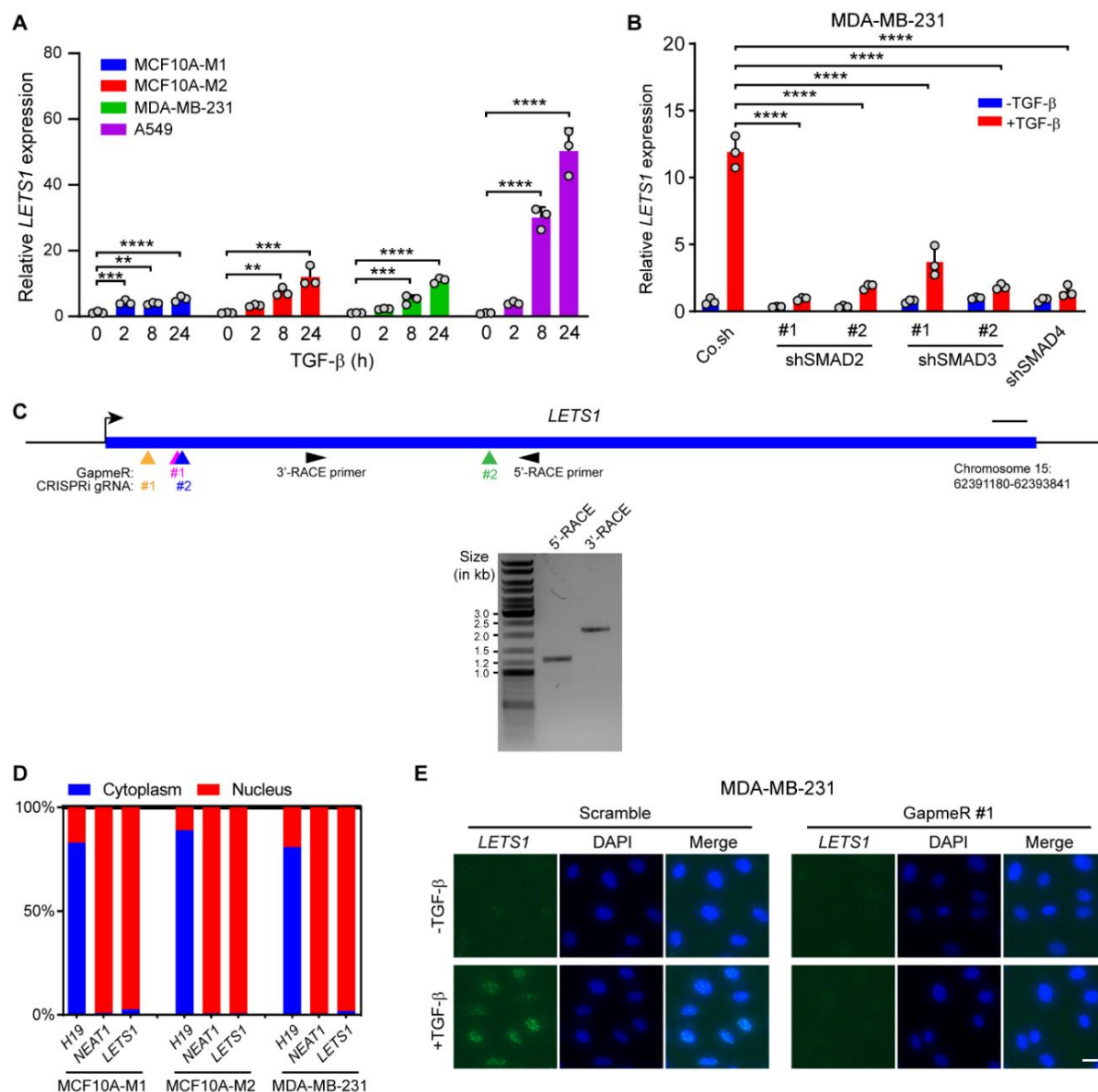


Fig. 1. *LETS1* is a nuclear lncRNA induced by TGF- β -SMAD signaling. (A) *LETS1* expression was measured by RT-qPCR in MCF10A-M1, MCF10A-M2, MDA-MB-231, and A549 cells. Cells were either not stimulated (0 hour) or stimulated with TGF- β for 2, 8, and 24 hours. (B) *LETS1* expression was measured by RT-qPCR in MDA-MB-231 cells upon shRNA-mediated *SMAD2*, *SMAD3*, or *SMAD4* knockdown. Co.sh, empty vector for shRNA expression. RT-qPCR results in (A) and (B) are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (C) Schematic illustration of the *LETS1* locus and the targeting regions of RACE primers, *LETS1*-targeting GapmeRs, and *LETS1*-targeting CRISPRi guide RNAs (gRNAs). Scale bar, 100 bps. *LETS1* 5' - and 3' -RACE DNA products were analyzed by agarose gel electrophoresis. (D) Subcellular distribution of lncRNAs *H19*, *NEAT1*, and *LETS1* based on RT-qPCR of the cytoplasmic and nuclear fractions of MCF10A-M1, MCF10A-M2, and MDA-MB-231 cells. Results are shown as means and representative of at least three independent experiments. The total amount of each lncRNA was set to 100%. (E) *LETS1* expression and subcellular localization was evaluated by RNA fluorescence in situ hybridization in MDA-MB-231 cells. Nuclei were stained with DAPI. Scale bar, 40 μ m. In (A) and (B), significance was assessed using one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparisons test. **, 0.001 < P < 0.01; ***, 0.0001 < P < 0.001; **** P < 0.0001.

In addition, *LETS1* knockdown alleviated TGF- β -induced filamentous (F)-actin formation in A549 cells (fig. S2C). The suppressive effect of *LETS1* knockdown on EMT was further confirmed by blocking *LETS1* transcription in MCF10A-M2 cells using CRISPR interference (CRISPRi) (fig. S2, D and E). In contrast, ectopic *LETS1* expression potentiated TGF- β -induced EMT marker expression in A549 cells (Fig. 2B and fig. S2, F and G). Transcriptional

profiling and gene set enrichment analysis (GSEA) also validated the positive correlation between the manipulation of *LETS1* expression and the EMT gene signature (Fig. 2C).

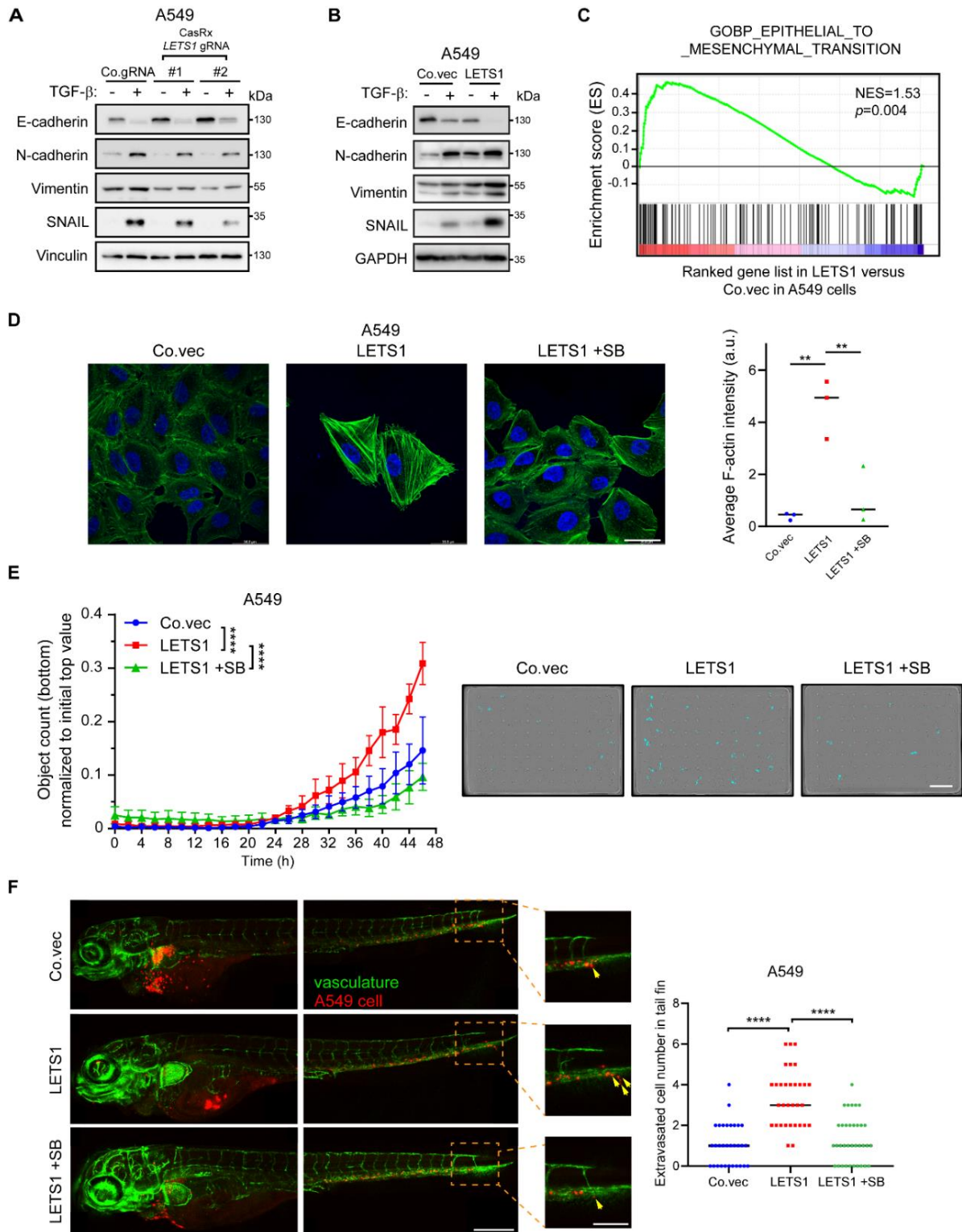


Fig. 2. *LETS1* promotes TGF- β -induced EMT, migration, and extravasation in cancer cells. (A and B) Immunoblotting for E-cadherin, N-cadherin, vimentin, and SNAIL in A549 cells expressing CRISPR-CasRx construct and empty vector (Co.gRNA) or *LETS1*-targeting gRNA (A) and in A549 cells overexpressing *LETS1* or empty vector (Co.vec) (B). Vinculin and GAPDH are loading controls. Blots are representative of at least three independent experiments. (C) GSEA of the correlation between experimentally manipulated *LETS1* expression and the EMT gene signature in A549 cells. NES, normalized enrichment score. (D) Fluorescent staining for F-actin in A549 cells overexpressing *LETS1* or empty vector

(Co.vec). DAPI staining was performed to visualize nuclei. Scale bar, 38.8 μm . Quantification of average F-actin intensity is shown as means \pm SD from three independent experiments. a.u., arbitrary units. (E) An IncuCyte chemotactic migration assay was performed with A549 cells overexpressing *LETS1* and treated with SB431542 (SB) or vehicle during the migration assays. Cells that migrated to the bottom chambers are marked in blue in the images. The migration results are expressed as means \pm SD from four biological replicates in one independent experiment and representative of at least three independent experiments. Scale bar, 400 μm . (F) *In vivo* zebrafish extravasation experiments with A549 cells stably expressing mCherry (red) and the *LETS1* expression construct or empty vector (Co.vec). A549 cells were injected into zebrafish embryos expressing enhanced green fluorescent protein (EGFP) throughout the vasculature and treated with vehicle or SB431542 (SB). Extravasated lung cancer cells in the zoomed tail fin area are indicated with yellow arrows. Numbers of extravasated cells are expressed as means \pm SD. Scale bars, 309.1 (whole fish) and 154.5 μm (enlargements). N = at least 30 fish per treatment group. In (C), significance was assessed by Kolmogorov-Smirnov test. In (D) and (F), significance was assessed using one-way ANOVA followed by Dunnett's multiple comparisons test. In (E), significance was assessed using two-way ANOVA followed by Tukey's multiple comparisons test. **, 0.001 < P < 0.01; **** P < 0.0001.

To test the effect of *LETS1* on cell migration, we performed chemotactic migration assays in MDA-MB-231 cells. As expected, CRISPRi-mediated *LETS1* knockdown alleviated TGF- β -induced cell migration (fig. S3A). In agreement with this result, *LETS1* depletion resulted in a decrease of MDA-MB-231 cell extravasation in a zebrafish xenograft cancer model (fig. S3B). On the contrary, *LETS1* ectopic expression promoted F-actin formation, migration, and extravasation in A549 cells (Fig. 2, D to F). Of note, TGF- β signaling blockage using the selective T β RI kinase inhibitor SB431542 (SB) mitigated the tumor-promoting effect of *LETS1* overexpression on A549 cells (Fig. 2, D to F). These findings indicate that *LETS1* is a pivotal potentiator of TGF- β -induced EMT, migration, and extravasation in lung and breast cancer cells.

***LETS1* potentiates TGF- β -SMAD signaling**

Because the results above suggested that *LETS1* may act as a modulator of TGF- β signaling, we investigated the effect of *LETS1* on TGF- β -SMAD signaling transduction. We observed that CRISPRi-mediated *LETS1* knockdown reduced, whereas *LETS1* ectopic expression enhanced, the activity of a highly selective synthetic reporter of transcription driven by SMAD3 and SMAD4 (SMAD3/4)³⁸ in HepG2 cells (Fig. 3A). Consistently, ectopic *LETS1* expression potentiated transcriptional activity of a SMAD3/4-driven dynamic green fluorescent protein (GFP) reporter (Fig. 3B)³⁹. Moreover, GapmeR-mediated *LETS1* knockdown suppressed the expression of TGF- β -induced target genes (*PAI-1*, *CTGF*, and *SMAD7*) in MDA-MB-231 and A549 cells (Fig. 3C and fig. S4A). However, ectopic *LETS1* expression promoted the TGF- β -SMAD-induced transcriptional events, as shown by the increase in TGF- β target gene expression and the positive correlation between manipulated *LETS1* expression and the TGF- β gene response signature⁴⁰ in A549 cells (Fig. 3, D and E). Furthermore, *LETS1* knockdown decreased, whereas *LETS1* overexpression increased, the TGF- β -induced SMAD2 phosphorylation in MDA-MB-231, A549, and MCF10A-M2 cells (Fig. 3, F and G; and fig. S4, B and C).

***LETS1* inhibits T β RI polyubiquitination and promotes T β RI stability by inducing *NR4A1* expression**

Given that *LETS1* potentiated TGF- β signaling upstream of SMAD2 phosphorylation, we tested the effect of *LETS1* on T β RI, the TGF- β receptor that directly mediates SMAD2/3 activation. Although the total T β RI protein abundance remained unaffected, the amount of T β RI at the plasma membrane was significantly reduced in the absence of *LETS1* in MDA-MB-231 cells (Fig. 4A and fig. S5A). Consistent with this notion, we found that T β RI polyubiquitination was increased upon *LETS1* knockdown, whereas *LETS1* overexpression reduced T β RI polyubiquitination in MDA-MB-231 cells (Fig. 4B and fig. S5B). Considering the nuclear localization of *LETS1*, we hypothesized that the transcription of TGF- β -SMAD signaling modulators may be altered by *LETS1*. To screen for relevant *LETS1* target genes, we

The lncRNA *LETS1* promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism

analyzed the changes in the transcriptional profile of A549 cells upon ectopic *LETS1* expression. As expected, transcripts of multiple TGF- β target genes, including *FOSB*, *COL11A1*, *JUN*, *JUNB*, *ATF3*, and *SNAIL*, were significantly increased by ectopic *LETS1* expression (Fig. 4C).

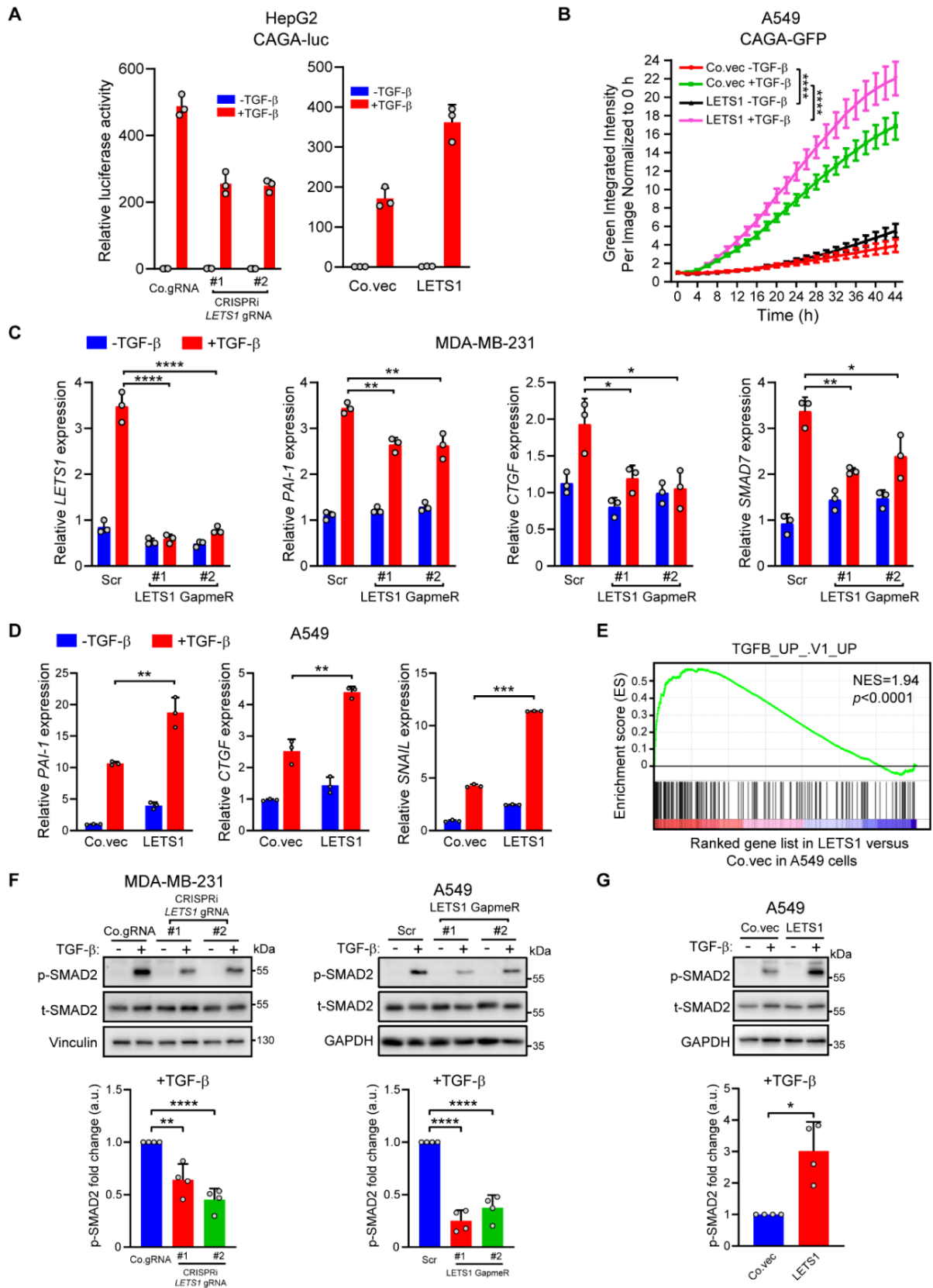


Fig. 3. *LETSI* potentiates TGF- β /SMAD signaling. (A) Quantification of luciferase activity in HepG2 cells expressing the synthetic SMAD3/4-responsive reporter CAGA-luc and either the *LETSI*-targeting CRISPRi gRNA construct or the *LETSI* overexpression construct and stimulated with TGF- β or vehicle. Co.gRNA and Co.vec are the corresponding empty vectors. The relative luciferase activities are representative of at least three independent experiments and expressed as means \pm SD from three wells of cells per treatment group in one experiment. (B) Quantification of GFP fluorescence in A549 cells coexpressing the CAGA-GFP reporter and either empty vector (Co.vec) or *LETSI* overexpression construct and stimulated with TGF- β or vehicle. The results are expressed as means \pm SD from six biological replicates in one independent experiment and representative of two independent experiments. (C) Quantification of *LETSI*, *PAI-1*, *CTGF*, and *SMAD7* expression in MDA-MB-231 cells transfected with GapmeRs targeting *LETSI* and treated with TGF- β or vehicle. Scr, scrambled GapmeR. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (D) Quantification of *PAI-1*, *CTGF*, and *SNAIL* expression in A549 cells overexpressing *LETSI* or empty vector and treated with TGF- β or vehicle. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (E) GSEA of correlation between experimentally manipulated *LETSI* expression and the TGF- β gene response signature in A549 cells. NES, normalized enrichment score. Significance was assessed by Kolmogorov-Smirnov test. (F and G) Immunoblotting for phosphorylated (p-) and total (t-)SMAD2 in TGF- β -stimulated MDA-MB-231 or A549 cells in which *LETSI* was knocked down by CRISPRi (MDA-MB-231) or GapmeR (F) or in which *LETSI* was overexpressed (G). Vinculin and GAPDH are loading controls. Quantitative data show the abundance of p-SMAD2 relative to t-SMAD2. Data are means \pm SD from four independent experiments. a.u., arbitrary units. In (B), significance was assessed using two-way ANOVA followed by Tukey's multiple comparisons test. In (C), significance was assessed using one-way ANOVA followed by Dunnett's multiple comparisons test. In (D), significance was assessed using unpaired Student's t test. In (E), significance was assessed by Kolmogorov-Smirnov test. In (F) and (G), significance was assessed using paired Student's t test. *, 0.01 < P < 0.05; **, 0.001 < P < 0.01; ***, 0.0001 < P < 0.001; **** P < 0.0001.

Furthermore, we found that *LETSI* promoted the expression of transcripts encoding NR4A1, which potentiates TGF- β -SMAD signaling by inhibiting T β RI polyubiquitination in breast cancer cells (Fig. 4, C to E; and fig. S5, C and D)²⁰. *Cis*-regulation is a mechanism by which nuclear lncRNAs can affect the expression of neighboring genes⁴¹. However, expression of genes near *LETSI* was not affected by ectopic *LETSI* expression in A549 cells (fig. S8B). This excludes the involvement of *LETSI* in a *cis*-regulatory mechanism.

We next determined whether *LETSI* exerted its function by inducing *NR4A1* expression. Upon *NR4A1* depletion by a selective shRNA or a mixture of four siRNAs, the promotion of TGF- β -SMAD3-driven transcriptional response induced by *LETSI* was alleviated in HepG2 cells (Fig. 4F and fig. S5E). Moreover, we demonstrated that *NR4A1* depletion attenuated *LETSI*-mediated promotion of EMT marker expression and migration in A549 cells (Fig. 4, G and H; and fig. S5, F to L). Together, our results suggest that *LETSI* induces *NR4A1* expression to suppress T β RI polyubiquitination and enhance TGF- β -SMAD signaling, EMT, and migration in cancer cells.

NFAT5 interacts with *LETSI*, inhibits T β RI polyubiquitination, and potentiates TGF- β -induced EMT and cell migration

To determine whether *LETSI* affected *NR4A1* expression at the transcriptional level, we cloned the 1597-base pair (bp) *NR4A1* promoter [P1; chromosome 12: 52,040,360 to 52,041,947 (GRCh38.p14)] and placed it upstream of a luciferase reporter gene (Fig. 5A). Ectopic *LETSI* expression enhanced transcriptional activity of the *NR4A1* P1 promoter, and further analysis of *NR4A1* promoter truncation mutants suggested that the promoter region containing bps -1238 to -1004 [chromosome 12: 52,040,567 to 52,040,801 (GRCh38.p14)] was required for *LETSI*-driven transcriptional activity (Fig. 5A). Nuclear lncRNAs can participate in gene transcription by interacting with transcription factors or chromatin modifiers^{21, 42}. We therefore applied the CRISPR-assisted RNA-protein interaction detection method (CARPID)⁴³ followed by mass spectrometry to identify nuclear protein partners of *LETSI* (fig. S6A). A well-characterized transcription factor, NFAT5, was enriched as one of the proteins with the highest binding capabilities to *LETSI* (Fig. 5B). We validated the *LETSI*-NFAT5 interaction in the presence or absence of TGF- β . Short TGF- β stimulation (1 hour) induced a moderate increase in *LETSI* expression (fig. S6B) but potently promoted *LETSI*-NFAT5 interaction (fig. S6C). Moreover,

the interaction between endogenous *LETS1* and endogenous NFAT5 was confirmed using RNA immunoprecipitation (RIP; Fig. 5C and fig. S6D) in MDA-MB-231 cells and between *in vitro*-transcribed *LETS1* and epitope-tagged NFAT5 using RNA pull-down assays in human embryonic kidney (HEK)293T cells (Fig. 5D).

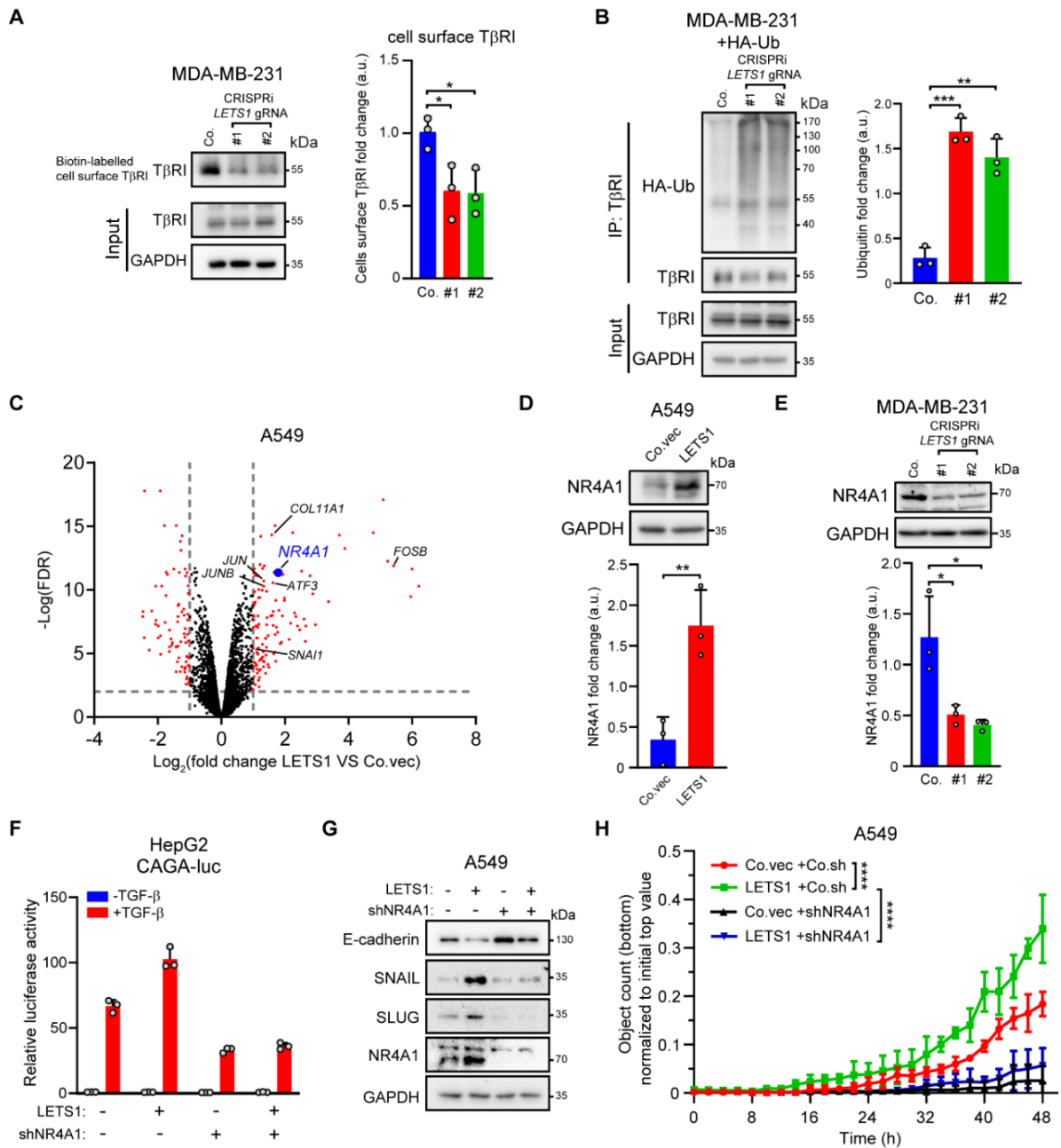


Fig. 4. *LETS1* inhibits T β RI polyubiquitination and promotes T β RI stability by inducing NR4A1 expression. (A) Immunoblotting and quantification of T β RI in total cell lysates (input) and biotinylated surface proteins from MDA-MB-231 cells in which *LETS1* was depleted by CRISPRi. Co, empty vector control. GAPDH is a loading control. Results are means \pm SD from three independent experiments. a.u., arbitrary units. (B) Immunoblotting for HA and T β RI in total lysates (input) and T β RI immunoprecipitates (IP) from MDA-MB-231 cells expressing HA-ubiquitin (HA-Ub) and empty vector or CRISPRi-gRNAs targeting *LETS1*. Ubiquitin was quantified in the T β RI immunoprecipitates. Quantitative data are means \pm SD from three independent experiments. (C) Volcano plot showing differentially expressed genes (as analyzed by RNA-seq) upon *LETS1* ectopic expression in A549 cells. (D and E) Immunoblotting and quantification of NR4A1 in A549 cells overexpressing *LETS1* (D) and in MDA-MB-231 cells in which *LETS1* was depleted by CRISPRi (E). Co. vec and Co., empty vector controls. Results are means \pm SD from three independent experiments. (F) Luciferase activity in TGF- β -stimulated HepG2 cells transfected with the expression construct for the SMAD3/4 transcriptional reporter CAGA-luc plus the *LETS1* ectopic expression construct and the NR4A1 shRNA construct as indicated. The relative luciferase activities are representative of at

least three independent experiments and expressed as means \pm SD from three wells of cells per treatment group in one experiment. (G) Immunoblotting for E-cadherin, SNAIL, SLUG, and NR4A1 in A549 cells in which LETS1 was overexpressed and NR4A1 was knocked down as indicated. Blots are representative of at least three independent experiments. (H) Quantification of migrated cells in IncuCyte chemotactic migration assays using A549 cells with LETS1 overexpression and NR4A1 knockdown as indicated. The results are expressed as means \pm SD from five biological replicates in one independent experiment and representative of three independent experiments. In (A), (B), (D), and (E), significance was assessed using paired Student's t test. In (H), significance was assessed using two-way ANOVA followed by Tukey's multiple comparisons test. *, $0.01 < P < 0.05$; **, $0.001 < P < 0.01$; ***, $0.0001 < P < 0.001$; **** $P < 0.0001$.

We next investigated the effect of NFAT5 on TGF- β -SMAD signaling. Ectopic NFAT5 expression enhanced the TGF- β -induced transcriptional response in MCF10A-M2 cells and SMAD2 phosphorylation in MDA-MB-231 cells (Fig. 5, E and F; and fig. S6, E to G). In samples of patients with breast cancer or lung adenocarcinoma, we observed strong positive correlations between *NFAT5* expression and the TGF- β gene response signature (fig. S6H). Moreover, *NFAT5* knockdown promoted T β RI polyubiquitination in MDA-MB-231 cells (Fig. 5G). Furthermore, NFAT5 enhanced TGF- β -induced EMT marker expression and cell migration in MCF10A-M2 cells (Fig. 5, H and I; and fig. S7, A to C). In addition, *NFAT5* expression and *NR4A1* expression showed a positive correlation with the EMT signature in tumor samples from cohorts of patients with breast cancer or lung adenocarcinoma, respectively (fig. S7D).

***LETS1* induces NR4A1 expression by cooperating with NFAT5**

Because *LETS1* interacts with NFAT5 and activates *NR4A1* transcription, we hypothesized that NFAT5 was likely to be involved in *LETS1*-induced *NR4A1* expression. As expected, ectopic NFAT5 expression increased *NR4A1* promoter reporter activity in HepG2 cells and *NR4A1* expression in MCF10A-M2 cells (Fig. 6, A to C). Moreover, positive correlations between *NFAT5* and *NR4A1* expression were observed in tumor samples from patients with breast cancer or lung adenocarcinoma (Fig. 6D). To further test whether NFAT5 was required for *LETS1*-mediated *NR4A1* expression, we knocked down *NFAT5* in HepG2 cells ectopically expressing *LETS1*. *LETS1*-induced *NR4A1* promoter activity was attenuated upon *NFAT5* depletion (Fig. 6E). Consistently, *LETS1*-induced *NR4A1* expression was also reduced in MDA-MB-231 cells in which *NFAT5* was knocked down (Fig. 6F). We then analyzed the identified *NR4A1* minimal promoter (P5) sequences and mapped two putative NFAT5-binding sites [chromosome 12: 52,040,615 to 52,040,632 (GRCh38.p14); fig. S8A]. Chromatin IP (ChIP) assays demonstrated strong NFAT5 binding to the *NR4A1* promoter in MDA-MB-231 cells, and ectopic expression of *LETS1* potentiated this (Fig. 6G), indicating that *LETS1* enhances the binding ability of NFAT5 to the *NR4A1* promoter.

Discussion

In this study, we showed that TGF- β -SMAD-induced nuclear *LETS1* associated with the transcription factor NFAT5 to facilitate the transcription of *NR4A1*. NR4A1 inhibits T β RI polyubiquitination and enhances T β RI stability by promoting SMAD7 protein degradation²⁰, resulting in an increase in TGF- β -SMAD signaling, TGF- β -induced EMT, and cancer cell migration and extravasation (Fig. 6H). Thus, we found a previously unidentified mechanism by which TGF- β -SMAD signaling is fine-tuned at the receptor level through a specific unannotated lncRNA, *LETS1*. This mechanism is distinct from previous reports of lncRNAs regulating *TBRI* mRNA expression at the transcriptional³⁰ or posttranscriptional^{44–52} level.

The lncRNA *LETS1* promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism

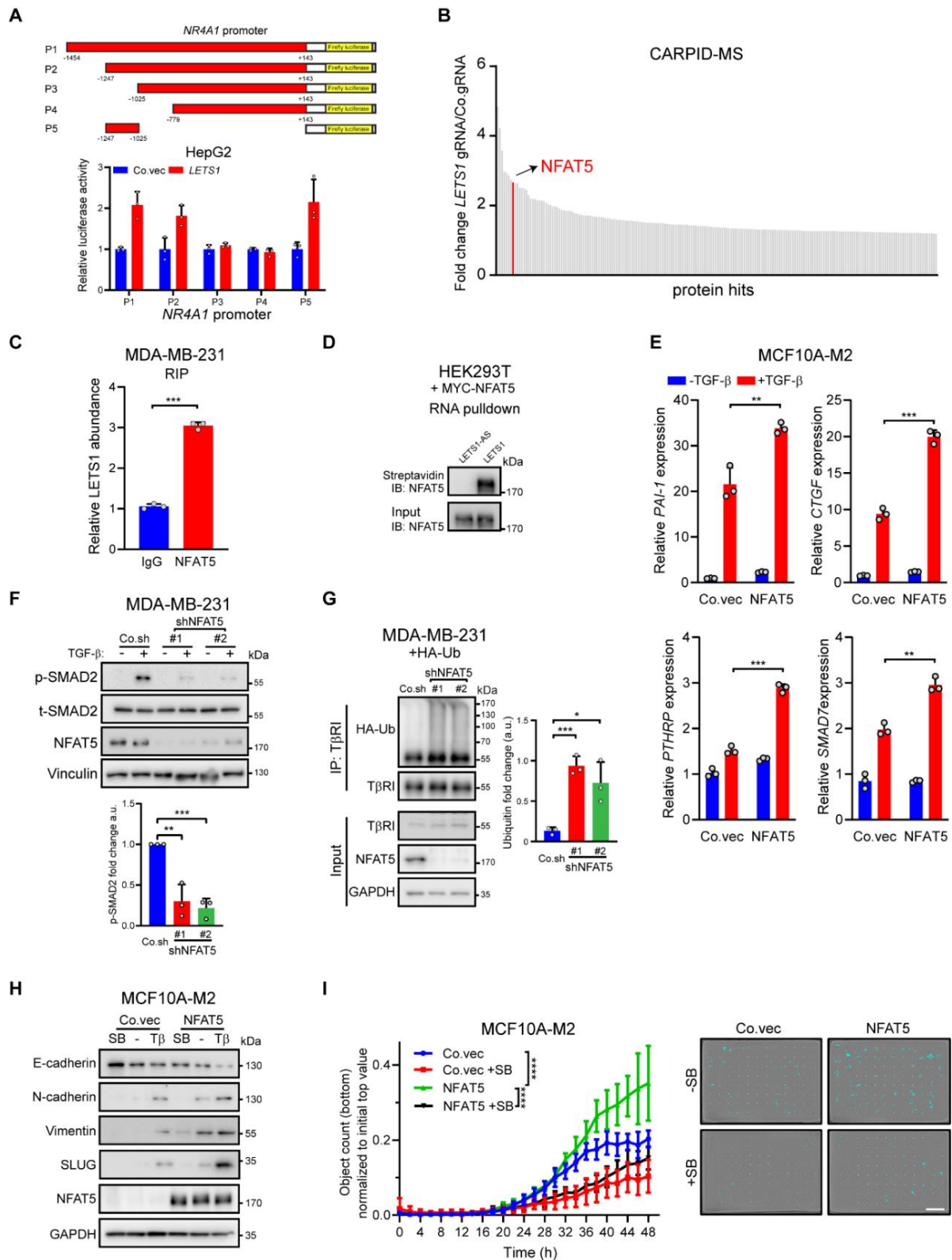


Fig. 5. NFAT5 interacts with *LETS1*; inhibits T β RI polyubiquitination; and potentiates TGF- β –SMAD signaling, EMT, and cell migration. (A) Quantification of luciferase activity in HEPG2 cells coexpressing the indicated *NR4A1* promoter luciferase reporter construct and the *LETS1* ectopic expression construct or empty vector (Co.vec). The relative luciferase activities are representative of at least three independent experiments and expressed as means \pm SD from three wells of cells per treatment group in one experiment. (B) Proteins that interact with *LETS1* were identified by CARPID followed by mass spectrometry (MS). The top 200 hits are shown, and the bar corresponding to *NFAT5* is indicated. (C) RIP assay quantifying *LETS1* abundance in *NFAT5* immunoprecipitates from MDA-MB-231 cells. *LETS1* abundance in *NFAT5* immunoprecipitates is presented as relative to that in IgG immunoprecipitates. RT-qPCR results are shown as means \pm SD from three biological

replicates in one independent experiment and representative of at least three independent experiments. (D) Immunoblotting (IB) for NFAT5 in total cell lysates (input) from HEK293T cells expressing MYC-NFAT5 and RNA pull-down assays in which the cell lysates were incubated with biotinylated antisense *LETS1* (*LETS1-AS*) or *LETS1* and affinity-purified with streptavidin beads. Blots are representative of at least three independent experiments. (E) Expression of *PAI-1*, *CTGF*, *PTHRP*, and *SMAD7* in MCF10A-M2 cells overexpressing NFAT5 and stimulated with TGF- β or vehicle. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (F) Immunoblotting for p-SMAD2 and t-SMAD2 and NFAT5 in TGF- β -stimulated MDA-MB-231 cells in which *NFAT5* was knocked down by two independent shRNAs. Quantitative data show the abundance of p-SMAD2 relative to t-SMAD2. Vinculin is a loading control. Results are means \pm SD from three independent experiments. a.u., arbitrary units. (G) Immunoblotting for HA and T β RI in total lysates (input) and T β RI immunoprecipitates (IP) from MDA-MB-231 cells expressing HA-Ub and transfected with empty vector (Co.sh) or *NFAT5*-targeting shRNA. Ubiquitin was quantified in the T β RI immunoprecipitates. GAPDH is a loading control. Results are means \pm SD from three independent experiments. (H) Immunoblotting for E-cadherin, N-cadherin, Vimentin, SLUG, and NFAT5 in MCF10A-M2 cells overexpressing NFAT5 or empty vector and treated with vehicle (-), SB431542 (SB), or TGF- β (T β). Blots are representative of at least three independent experiments. (I) Quantification of migrated cells in IncuCyte chemotactic migration assays using MCF10A-M2 cells overexpressing NFAT5 and treated with SB431542 or vehicle. The cells that migrated to the bottom chambers are marked in blue in the images. The migration results are expressed as means \pm SD from 12 biological replicates in one independent experiment and representative of at least three independent experiments. Scale bar, 400 μ m. In (C) and (E), significance was assessed using unpaired Student's t test. In (F) and (G), significance was assessed using paired Student's t test. In (I), significance was assessed using two-way ANOVA followed by Tukey's multiple comparisons test. *, 0.01 < P < 0.05; **, 0.001 < P < 0.01; ***, 0.0001 < P < 0.001; ****, P < 0.0001.

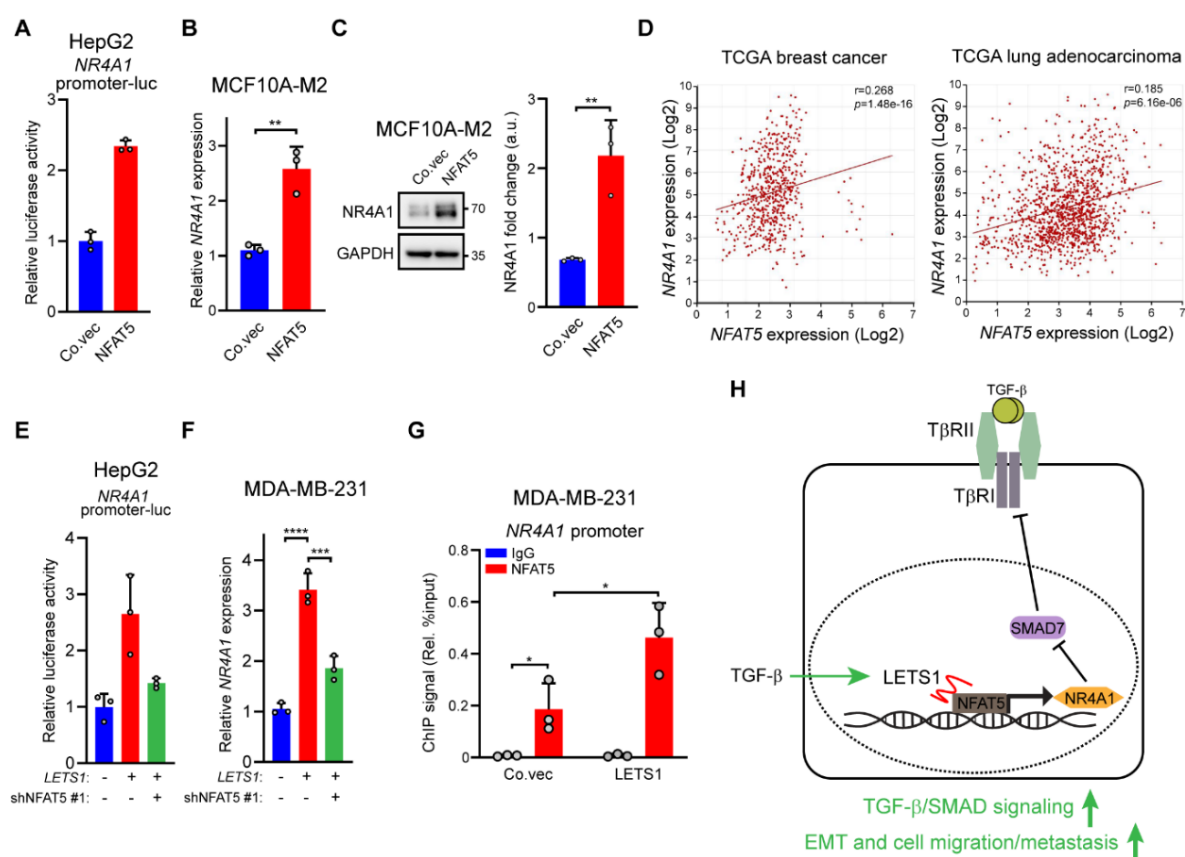


Fig. 6. *LETS1* cooperates with NFAT5 to induce *NR4A1* expression. (A) Quantification of luciferase activity in HepG2 cells coexpressing the *NR4A1* promoter luciferase reporter P5 and the NFAT5 expression construct or empty vector (Co.vec). The relative luciferase activities are representative of at least three independent experiments and expressed as means \pm SD from three wells of cells per treatment group in one experiment. (B) *NR4A1* expression in MCF10A-M2 cells transfected with the NFAT5 expression construct or empty vector. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (C) Immunoblotting for NR4A1 in MCF10A-M2 cells overexpressing NFAT5 or transfected with empty vector. GAPDH is a loading control. Results are means \pm SD from three independent experiments. a.u., arbitrary units. (D) Correlations between *NFAT5* and *NR4A1* expression in samples of patients with breast cancer or lung adenocarcinoma. (E) Quantification of *NR4A1* promoters luciferase reporter activity in HepG2 cells transfected with the *LETS1* expression construct and the shNFAT5 no. 1 construct as indicated. The relative luciferase activities are representative of at least three independent experiments and expressed as means \pm SD from three wells of cells per treatment group in one experiment. (F) Quantification of *NR4A1* expression in MDA-MB-231 cells expressing the *LETS1* expression construct and the shNFAT5 no. 1 construct as indicated. RT-qPCR results are shown as means

The lncRNA *LETS1* promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism

\pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (G) ChIP analysis of the *NR4A1* promoter region in MDA-MB-231 cells transduced with the *LETS1* expression construct or empty vector. IgG was included as the control for IP. RT-qPCR results are shown as means \pm SD from three independent experiments. (H) Schematic model of the action of *LETS1* on TGF- β -SMAD signal transduction through the potentiation of NFAT5-mediated *NR4A1* transcription. In (B), significance was assessed using unpaired Student's t test. In (C) and (G), significance was assessed using paired Student's t test. In (D), the statistical analysis was performed using Pearson's correlation (r) test. In (F), significance was assessed using one-way ANOVA followed by Dunnett's multiple comparisons test. *, $0.01 < P < 0.05$; **, $0.001 < P < 0.01$; ***, $0.0001 < P < 0.001$; **** $P < 0.0001$.

The pivotal promoting effects of *LETS1* on TGF- β -SMAD signaling and on TGF- β -induced EMT and migration were shown in our study by multiple orthogonal approaches, including GapmeRs, CRISPRi, CRISPR-CasRx, and ectopic expression to manipulate *LETS1* expression. Moreover, possible shortcomings with each approach were as much as possible controlled for. For example, off-target effects of *LETS1* targeting by CRISPRi⁵³ on neighboring gene expression were excluded (fig. S8B). Results of *LETS1* misexpression were shown in multiple cell lines, and *in vitro* cell culture studies were complemented with experiments using the *in vivo* zebrafish embryo xenograft model for extravasation. Conservation of the lncRNA sequence is much lower than that of protein-coding RNAs among vertebrates⁵⁴. However, lncRNA orthologs with similar secondary or tertiary structures but diverse sequences may exert the same functions in different species⁵⁵. We performed a sequence similarity search for *LETS1* in the mouse transcriptome, but no ortholog of *LETS1* was identified, making genetic analysis of *LETS1* function in mouse cancer models challenging.

Cell surface T β RI is highly dynamic and undergoes rapid degradation after being polyubiquitinated by E3 ligases such as SMURF2 and NEDD4^{56, 57}. As an adaptor of T β RI and E3 ligase interactions, SMAD7 potentiates the E3 ligase-mediated polyubiquitination of T β RI^{56, 57}. NR4A1 potentiates TGF- β -SMAD signaling by enhancing SMAD7 degradation in breast and lung cancer cells^{20, 58, 59}. Our results showed that *NR4A1* knockdown greatly mitigated the promoting effects of *LETS1* on TGF- β signaling, TGF- β -induced EMT, and cell migration, suggesting that NR4A1 is a major *LETS1* downstream effector. However, because the expression of multiple genes was altered upon ectopic *LETS1* expression in our transcriptome analysis, other genes also likely participate in the effects mediated by *LETS1*.

NFAT5 was identified as a protein partner of *LETS1*, and TGF- β stimulation potently promoted *LETS1*-NFAT5 interaction in MDA-MB-231 cells. A possible explanation for this result could be that TGF- β treatment alters the chemical modification (such as N6-methyladenosine) of *LETS1* and/or posttranslational modification (such as phosphorylation) of NFAT5, thereby promoting this interaction. Therefore, further investigation is required to explore these and other possibilities.

We showed that NFAT5 directly bound to the *NR4A1* promoter and stimulated its activity, which was strengthened upon *LETS1* ectopic expression. Previous reports have documented that the promoter activity of *NR4A1* can be enhanced by the transcription factor CCAAT/enhancer binding protein β (C/EBP β) and several lysine methyltransferases that are recruited by *LncLy6C*⁶⁰. Compared with other NFAT member proteins, NFAT5 lacks the structural domain that mediates the cooperative complex formation with other transcription factors^{61, 62}. It is possible that the interaction with *LETS1* may provide extra docking sites on NFAT5 for other proteins to potentiate NFAT5 transcriptional activity or for chromatin modifiers to change the local chromatin status. In addition, the C-terminal dimerization of NFAT5 is required for its DNA binding activity⁶³. *LETS1* may facilitate the formation of NFAT5 homodimers or stabilize the dimeric complex through its binding to NFAT5. Because the affinity of NFAT5 for DNA is much lower than that of other NFAT family members⁶¹,

another possibility is that the interaction with *LETS1* may change the conformation of NFAT5 toward a status with stronger DNA binding ability. However, whether the contribution of *LETS1* to NFAT5-mediated transcription is confined to a certain subset of target genes including *NR4A1* or this effect can be expanded to general transcriptional events directed by NFAT5 requires further investigation.

Our results showed that NFAT5 is a positive regulator of TGF- β -induced EMT and cell migration in breast and lung cancer cells. These results are consistent with other studies demonstrating the tumor-promoting role of NFAT5 through the induction of the expression of genes encoding proteins such as aquaporin-5 and S100 calcium binding protein A4^{64–67}. We found that TGF- β -SMAD signaling was required for NFAT5 to induce EMT and migration in cell culture models and observed strong correlations between *NFAT5* expression and the TGF- β response gene signature or the EMT signature in RNA profiles obtained from biopsies of patients with breast cancer or lung adenocarcinoma. These results reveal a previously undescribed mechanism by which NFAT5 promotes cancer progression and highlight the therapeutic potential of targeting NFAT5 in cancer. Compared with enzymes and kinases, transcription factors are difficult to target with small-molecule inhibitors because of the lack of active sites or allosteric regulatory pockets⁶⁸. DNA-based proteolysis targeting chimera (PROTAC) approaches such as transcription factor (TF)-PROTAC⁶⁹ and oligonucleotide-based PROTAC⁷⁰ have been developed to selectively and efficiently degrade transcription factors of interest. Therefore, on the basis of the consensus DNA binding sequence of NFAT5, NFAT5-specific DNA oligomers could be designed and combined with the E3 ligase ligands typically used for TF-PROTAC to target NFAT5 for degradation in cancer cells.

In conclusion, we identified *LETS1* as a potent activator of TGF- β -induced EMT and cancer cell migration and extravasation, all of which contribute to cancer progression, by promoting T β RI cell surface abundance. Inhibition of *LETS1* expression, for example, using GapmeR⁷¹ or ribonuclease-targeting chimera (RIBOTAC)⁷² approaches, may therefore have therapeutic potential in cancer.

Materials and methods

Cell culture and reagents

HEK293T (CRL-1573), HepG2 (HB-8065), A549 (CRM-CCL-185), and MDA-MB-231 (CRM-HTB-26) cells were purchased from the American Type Culture Collection. MCF10A-M1 and MCF10A-M2 cells were provided by F. Miller (Barbara Ann Karmanos Cancer Institute, Detroit, USA). All the cell lines were cultured as described previously⁷³. Recombinant TGF- β 3 was a gift from A. P. Hinck (University of Pittsburgh). Cells were frequently tested for absence of mycoplasma, and cell lines were authenticated by short tandem repeat profiling.

Plasmid construction

LETS1 cDNA was cloned from A549 cells and ligated to the pCDH-EF1 α -MCS-polyA-PURO lentiviral vector. Guide RNAs (gRNAs) for CRISPRi and CRISPR-CasRx were inserted into the pLKO.1-U6-PURO (AA19) and pRX004-pregRNA (Addgene, 109054), respectively. *NR4A1* promoter fragments were cloned into the pGL4-luc backbone (Promega). The primers used for molecular cloning are listed in table S1.

Lentiviral transduction and transfection

Production of lentivirus was described elsewhere⁷³. Cells stably expressing the indicated constructs were selected by adding the corresponding antibiotics to the culture medium after 2

days postinfection. We used TRCN0000010477 (no. 1) and TRCN0000010478 (no. 2) for *SMAD2* knockdown, TRCN0000330128 (no. 1) and TRCN0000330127 (no. 2) for *SMAD3* knockdown, TRCN0000040031 for *SMAD4* knockdown, TRCN0000019426 for *NR4A1* knockdown, and TRCN0000020019 (no. 1) and TRCN0000020021 (no. 2) for *NFAT5* knockdown. For the transfection of GapmeRs (Eurogentec) and *NR4A1*-targeting SMARTpool siRNA (Horizon, L-003426), 1.2×10^5 A549 cells were seeded in wells of a 12-well plate and incubated with complex formed by Lipofectamine 3000 (Thermo Fisher Scientific, L3000015) and GapmeRs (25 nM) or siRNA (25 nM). Knockdown efficiency was quantified at 2 days after transfection. The sequences of GapmeRs are listed in table S2.

RT-qPCR

To check *LETS1* expression upon TGF- β stimulation, cells were starved for 16 hours and treated with vehicle control or TGF- β (5 ng/ μ l) for indicated durations as indicated in the panels or 4 hours, if the treatment duration is not specified. CHX (50 μ g/ml) was used to pretreat MDA-MB-231 cells for 30 min before adding TGF- β or vehicle. To evaluate TGF- β -induced target gene expression, cells were starved for 16 hours and treated with vehicle control or TGF- β (1 ng/ μ l) for 4 hours. RNA extraction and RT-qPCR were performed as described previously⁷³. Expression of target genes was normalized to GAPDH. The primer sequences used for RT-qPCR are listed in table S3.

Western blotting

To detect EMT marker expression, A549 or MCF10A-M2 cells were treated with TGF- β (1 ng/ml for A549 and 5 ng/ml for MCF10A-M2, respectively) or vehicle for 1 (A549) or 3 days (MCF10A-M2). To check TGF- β -induced p-SMAD2, TGF- β (1 ng/ml) or vehicle was added for indicated time points or 1 hour, if the treatment duration is not specified. Western blotting was performed as described previously⁷³. The primary antibodies are listed in table S4.

Coding potential prediction

CPAT software was used to predict the coding potential of protein-coding mRNAs or lncRNAs as described elsewhere³⁷.

Transcriptional reporter assays

Reporter assays were performed as described previously⁷³ to quantify SMAD3/4-driven transcriptional CAGA-luc reporter activity in HepG2 cells. Cells were serum-starved for 6 hours and stimulated with TGF- β (1 ng/ml) or vehicle control for 16 hours. To measure *NR4A1* promoter fragment activity, 320 ng of the *LETS1* or *NFAT5* expression construct, 100 ng of the *NR4A1* promoter luciferase reporter, and 80 ng of the β -galactosidase expression construct were cotransfected into HepG2 cells using polyethyleneimine (Polysciences, 23966). Luciferase activity was measured with the substrate d-luciferin (Promega) and a luminometer (PerkinElmer) and normalized to β -galactosidase activity.

Fluorescent staining

To evaluate the expression and localization of F-actin, fluorescent staining was performed as previously described^{74, 75}. Briefly, A549 cells were stimulated with SB431542 (SB; 10 μ M) or TGF- β (1 ng/ml) or the corresponding vehicle for 48 hours. The fixed cells were stained with phalloidin conjugated with Alexa Fluor 488 (1:500 dilution; Thermo Fisher Scientific, A12379) for 30 min at room temperature. VECTASHIELD Antifade Mounting Medium with DAPI (4',6-diamidino-2-phenylindole; Vector Laboratories, H-1200) was used to mount coverslips. A Leica SP8 confocal microscope (Leica Microsystems) was used to acquire images. Quantification of average F-actin intensity was performed using the ImageJ software.

Ubiquitination assay

Ubiquitination assay was performed as previously described⁷³ in MDA-MB-231 cells stably expressing hemagglutinin (HA)-ubiquitin.

Chemotactic migration and live-cell imaging using IncuCyte

An IncuCyte live-cell imaging system (Essen BioScience) was used to monitor cell chemotactic migration as previously described⁷³. Cells were treated with TGF- β (5 ng/ml) or vehicle during the assay. To quantify the dynamic GFP signal in A549 cells, 5×10^3 A549 cells with SMAD3/4-driven GFP expression³⁹ were seeded in a 96-well plate. Cells were serum-starved for 16 hours and stimulated with TGF- β (1 ng/ml) or vehicle, and the real-time green integrated intensity was monitored using the IncuCyte system³⁹.

Subcellular fractionation

In brief, cell pellets were lysed in buffer A [50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1% NP-40, and 0.25% sodium deoxycholate] for 15 min on ice. The supernatant was collected as the cytoplasmic fraction after centrifugation at 3000g for 5 min. Phosphate-buffered saline (PBS) was used to wash the pellet, which was then resuspended in buffer B [50 mM Tris-HCl (pH 7.4), 400 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, and 1% SD]. The supernatant was collected as the nuclear fraction after 20 min of incubation on ice and centrifugation at 12,000g for 15 min.

RACE

RACE was carried out on A549 cells using a SMARTer RACE 5' /3' Kit (TaKaRa, 634859). 5' /3' RACE products were cloned and transformed into competent cells, and 20 independent colonies were picked for Sanger sequencing.

RIP

RIP was performed using a Magna RIP RNA-Binding Protein IP Kit (Merck Millipore, 17-700). A total of 2.5 μ g of an anti-NFAT5 antibody (Thermo Fisher Scientific, PA1-023) or normal rabbit immunoglobulin G (IgG) were added to the cell lysates. To lower the background, we optimized the supplied instructions by adding a bead-blocking step. The magna beads were blocked with 5 μ l of yeast tRNA (Invitrogen, AM7119) and 5 μ l of bovine serum albumin (Invitrogen, AM2618) for 2 hours at 4°C before being used for IP.

RNA pull-down assay

A MEGAscript Kit (Thermo Fisher Scientific, AM1334) was used to in vitro transcribe antisense and sense *LETS1*, which were then biotinylated with an RNA 3' End Desthiobiotinylation Kit (Thermo Fisher Scientific, 20160). RNA pull-down assays were performed using a Magnetic RNA-Protein Pull-Down Kit (Thermo Fisher Scientific, 20164). NFAT5 expression was analyzed by Western blotting.

ChIP assay

Briefly, 1×10^7 MDA-MB-231 cells were cross-linked with 1% formaldehyde for 10 min and resuspended in lysis buffer [5 mM Pipes (pH 8.0), 85 mM KCl, and 0.5% NP-40] for 10 min on ice. After centrifugation at 500g for 5 min at 4°C, the pellet was lysed in nuclear lysis buffer [50 mM Tris-HCl (pH 8), 10 mM EDTA, and 1% SD] for 10 min on ice. Afterwards, the chromatin was sheared using a sonicator (Diagenode) at 30% amplitude for 3 min. After centrifugation at 12,000g for 30 min at 4°C, the supernatant was diluted five times with IP dilution buffer [50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% NP-40, and 0.25% sodium deoxycholate]. Protein A Sepharose beads (GE Healthcare, catalog no. 17-0963-03)

and the salmon sperm DNA were used to preclear the cell lysates for 1 hour at 4°C. Subsequently, the cell lysates were incubated with 10 μ g of IgG (Cell Signaling Technology, 2729) or anti-NFAT5 antibody (Thermo Fisher Scientific, PA1-023) overnight at 4°C. The next day, 20 μ g of Protein A Sepharose beads were added to the cell lysates and incubated for 2 hours at 4°C. After five times washing, the beads were treated with ribonuclease A and proteinase K, and the DNA was extracted by isopropanol. The amount of precipitated *NR4A1* promoter region was analyzed by RT-qPCR and the absolute quantification method.

CARPID and mass spectrometry

MDA-MB-231 cells stably expressing TurboID-dCasRx and CRISPR-CasRx gRNA was treated with TGF- β (2.5 ng/ml) or vehicle for 1 hour. Two hundred μ M biotin (Sigma-Aldrich, B4639) dissolved in medium was used to activate biotinylation in cells cultured in a 15-cm dish for 30 min. Cells were washed with cold PBS twice and suspended with 600 μ l of lysis buffer [50 mM Tris-HCl (pH 7.4), 500 mM NaCl, 0.4% SD, 5 mM EDTA, H₂O, and 1 mM dithiothreitol]. After mixing with 240 μ l of 20% Triton X-100, cell lysates were sonicated at 80% amplitude for 10 s four times. The supernatant was collected after centrifugation at 12,000g for 30 min at 4°C and added with 1 ml of 50 mM Tris-HCl (pH 7.5). Twenty-five microliters of Streptavidin Agarose beads (Millipore, 69203) were added to the supernatant and incubated on a rotator overnight at 4°C. After washing with wash buffer 1 (2% SD), wash buffer 2 [0.1% deoxycholate, 1% Triton X-100, 500 mM NaCl, 1 mM EDTA, and 50 mM Hepes (pH 7.5)], wash buffer 3 [250 mM LiCl, 0.5% Triton X-100, 0.5% deoxycholate, 1 mM EDTA, and 10 mM Tris-HCl (pH 8.1)], wash buffer 4 [50 mM Tris-HCl (pH 7.4) and 50 mM NaCl], and 50 mM ammonium bicarbonate three times, the beads were boiled for 5 min in sample buffer, and biotinylated proteins of interest were analyzed by Western blotting. For mass spectrometry analysis, the beads were resuspended in 250 μ l of 50 mM ammonium bicarbonate and incubated with 250 ng of trypsin (Promega, V5280) overnight at 37°C. The beads were separated with a prewashed 0.4- μ m filter (Millipore, UFC30HV00). Digested peptides were desalted using StageTips⁷⁶ and analyzed as in⁷⁷. Briefly, samples were measured in an Orbitrap Exploris 480 (Thermo Fisher Scientific) mass spectrometer coupled to an Ultimate 3000 Ultra-High-Performance Liquid Chromatography (Dionex). Digested peptides were separated using a 50-cm-long fused silica emitter (FS360-75-15-N-5-C50, New Objective, MA, USA) in-house packed with 1.9- μ m C18-AQ beads (Reprospher-DE, Pur, Dr. Maisch, Ammerburch-Entringen, Germany) and heated to 50°C in a Column Oven for electrospray ionization/Nanospray (Sonation, Germany). Peptides were separated by liquid chromatography using a gradient from 2 to 32% acetonitrile with 0.1% formic acid for 60 min, followed by column reconditioning for 25 min. A lock mass of 445.12003 (polysiloxane) was used for internal calibration. Data were acquired in a data-dependent acquisition mode with a TopSpeed method with cycle time of 3 s with a scan range of 350 to 1600 mass/charge ratio (m/z) and resolutions of 60,000 and 30,000 for MS1 and MS2, respectively. For MS2, an isolation window of 1.2 m/z and a higher-energy C-trap dissociation (HCD) collision energy of 30% were applied. Precursors with a charge of 1 and higher than 6 were excluded from triggering MS2 as well as previously analyzed precursors with a dynamic exclusion window of 30 s.

Mass spectrometry data analysis

Mass spectrometry data were analyzed using MaxQuant v2.1.3.0 according to Tyanova et al.⁷⁸ with the following modifications: Maximum missed cleavages by trypsin was set to 3. Searches were performed against an in silico-digested database from the human proteome including isoforms and canonical proteins (UniProt, 29 August 2022). Oxidation (M), acetyl (protein N-terminal), were set as variable modifications with a maximum of 3. Carbamidomethyl (C) was disabled as a fixed modification. Label-free quantification was activated not enabling fast label-

free quantification (LFQ). The match between runs feature was activated with default parameters.

MaxQuant output data were further processed in the Perseus Computational Platform v1.6.14.0 according to Tyanova et al.⁷⁹. LFQ intensity values were log₂-transformed, and potential contaminants and proteins identified by site only or reverse peptide were removed. Samples were grouped in experimental categories, and proteins not identified in three of three replicates in at least one group were also removed. Missing values were imputed using normally distributed values with a 2.1 downshift (log₂) and a randomized 0.1 width (log₂) considering whole-matrix values. Two-sided t tests were performed to compare groups. Analyzed data were exported from Perseus and further processed in Microsoft Excel 365 for comprehensive visualization. Protein hits were ranked on the basis of the fold change between two *LETS1*-targeting gRNAs and the control gRNA expression vector (Co.gRNA).

Transcriptional profiling and GSEA

To identify TGF- β -induced lncRNAs, cells were serum-starved overnight and stimulated without (0 hours) or with TGF- β (5 ng/ml) for 2, 8, and 24 hours. RNA was extracted using TRIzol reagent (Thermo Fisher Scientific, 15596026). Libraries were then constructed, and RNA sequencing (RNA-seq) was performed on an Illumina HiSeq [Beijing Genomics Institute (BGI), Shenzhen]. Differentially expressed lncRNAs were analyzed by BGI. To identify mRNAs affected by *LETS1*, we generated A549 cells stably expressing *LETS1*. The DNBSeg platform (BGI, Hong Kong) was used to perform RNA-seq. Analysis of differentially expressed genes was performed as described previously⁷³. The correlations between *LETS1* and TGF- β /SMAD signaling and EMT were performed with the GSEA software⁸⁰ using the TGF- β (TGFB_UP.V1_UP) gene response signature⁴⁰ and the EMT (GOBP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION; Gene Ontology: 0001837) gene signature as inputs.

Gene correlation analysis in databases

Correlations between *NFAT5* and *NR4A1* expression or between *NFAT5* expression and the TGF- β gene response signature or the EMT gene signature were performed in the breast (R2 internal identifier: ps_avgpres_tcgabrcav32a1221_gencode36) and lung (R2 internal identifier: ps_avgpres_tcgaluadv32a589_gencode36) cohorts of patients with cancer in the R2: Genomics Analysis and Visualization Platform (<http://r2.amc.nl>).

In situ hybridization staining

MDA-MB-231 cells were transfected with a scrambled GapmeR or *LETS1*-targeting GapmeR no. 1 and stimulated with TGF- β (5 ng/ml) or vehicle for 2 hours. The expression and localization of *LETS1* were detected by an RNAScope Multiplex Fluorescent kit (Advanced Cell Diagnostics, 323100) and an *in situ* probe for *LETS1* (Advanced Cell Diagnostics, 840831). A DMi8 inverted fluorescence microscope (Leica) was used to acquire images.

Embryonic zebrafish cancer cell extravasation assay

The experiments were conducted in a licensed establishment for the breeding and use of experimental animals [Leiden University (LU)] and subject to internal regulations and guidelines, stating that advice was taken from the Animal Welfare Body to minimize suffering for all experimental animals housed at the facility. The zebrafish assays described are not considered as an animal experiment under the Experiments on Animals Act (Wod, effective 2014), the applicable legislation in the Netherlands in accordance with the European guidelines (EU directive no. 2010/63/EU) regarding the protection of animals used for scientific purposes. Therefore, a license specific for these assays on zebrafish larvae (<5 days) was not required.

MDA-MB-231 or A549 cells labeled with mCherry were injected into the ducts of Cuvier of embryos from transgenic zebrafish [fli; enhanced GFP (EGFP)] as previously described⁸¹. Zebrafish embryos were maintained in 33°C egg water for 5 days. To check the effect of TGF- β signaling blockage on cell extravasation, SB431542 (SB; 1 μ M) or vehicle was added to egg water during the assay. Zebrafish were fixed with 4% formaldehyde. An inverted SP5 stimulated emission depletion (STED) confocal microscope (Leica) was used to visualize zebrafish embryos and injected cancer cells. At least 30 embryos per group were quantified. Two independent experiments were performed, and representative results are shown.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 9.3.1. All measurements in this study were taken from distinct samples.

References

1. Pastushenko, I. & Blanpain, C. EMT transition states during tumor progression and metastasis. *Trends Cell Biol.* **29**, 212-226 (2019).
2. Gui, P.L. & Bivona, T.G. Evolution of metastasis: new tools and insights. *Trends Cancer* **8**, 98-109 (2022).
3. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* **12**, 31-46 (2022).
4. Yang, J. *et al.* Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **21**, 341-352 (2020).
5. Sha, Y.T. *et al.* Intermediate cell states in epithelial-to-mesenchymal transition. *Phys. Biol.* **16** (2019).
6. Hendrix, M.J.C., Seftor, E.A., Seftor, R.E.B. & Trevor, K.T. Experimental co-expression of vimentin and keratin intermediate filaments in human breast cancer cells results in phenotypic interconversion and increased invasive behavior. *Am. J. Pathol.* **150**, 483-495 (1997).
7. Schliekelman, M.J. *et al.* Molecular portraits of epithelial, mesenchymal, and hybrid states in lung adenocarcinoma and their relevance to survival. *Cancer Res.* **75**, 1789-1800 (2015).
8. Dmello, C. *et al.* Vimentin regulates differentiation switch via modulation of keratin 14 levels and their expression together correlates with poor prognosis in oral cancer patients. *Plos One* **12**, e0172559 (2017).
9. George, J.T., Jolly, M.K., Xu, S., Somarelli, J.A. & Levine, H. Survival outcomes in cancer patients predicted by a partial EMT gene expression scoring metric. *Cancer Res.* **77**, 6415-6428 (2017).
10. Grosse-Wilde, A. *et al.* Stemness of the hybrid epithelial/mesenchymal state in breast cancer and its association with poor survival. *Plos One* **10**, e0126522 (2015).
11. Hao, Y., Baker, D. & ten Dijke, P. TGF- β -mediated epithelial-mesenchymal transition and cancer metastasis. *Int. J. Mol. Sci.* **20**, 2767 (2019).
12. Fan, C., Zhang, J., Hua, W. & ten Dijke, P. Biphasic role of TGF- β in cancer progression: From tumor suppressor to tumor promotor. *Reference Module in Biomed. Sci.*, (2018).
13. Tzavlaki, K. & Moustakas, A. TGF- β signaling. *Biomolecules* **10**, 487 (2020).
14. Hata, A. & Chen, Y.G. TGF- β signaling from receptors to SMADs. *Cold Spring Harb. Perspect. Biol.* **8**, a022061 (2016).
15. Yan, X., Xiong, X. & Chen, Y.G. Feedback regulation of TGF- β signaling. *Acta Biochim. Biophys. Sin.* **50**, 37-50 (2018).
16. Budi, E.H., Duan, D. & Derynck, R. Transforming growth factor- β receptors and SMADs: Regulatory complexity and functional versatility. *Trends Cell Biol.* **27**, 658-672 (2017).
17. Kavsak, P. *et al.* SMAD7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF β receptor for degradation. *Mol. Cell* **6**, 1365-1375 (2000).
18. Koinuma, D. *et al.* Arkadia amplifies TGF- β superfamily signalling through degradation of SMAD7. *EMBO J.* **22**, 6458-6470 (2003).
19. Zhang, L. *et al.* RNF12 controls embryonic stem cell fate and morphogenesis in zebrafish embryos by targeting SMAD7 for degradation. *Mol. Cell* **46**, 650-661 (2012).
20. Zhou, F. *et al.* Nuclear receptor NR4A1 promotes breast cancer invasion and metastasis by activating TGF- β signalling. *Nat. Commun.* **5**, 3388 (2014).
21. Statello, L., Guo, C.J., Chen, L.L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96-118 (2021).
22. Nandwani, A., Rathore, S. & Datta, M. LncRNAs in cancer: Regulatory and therapeutic implications. *Cancer Lett.* **501**, 162-171 (2021).
23. Lin, C. & Yang, L. Long noncoding RNA in cancer: Wiring signaling circuitry. *Trends Cell Biol.* **28**, 287-301 (2018).

24. Mattick, J.S. & Rinn, J.L. Discovery and annotation of long noncoding RNAs. *Nat. Struct. Mol. Biol.* **22**, 5–7 (2015).
25. Palazzo, A.F. & Koonin, E.V. Functional long non-coding RNAs evolve from junk transcripts. *Cell* **183**, 1151–1161 (2020).
26. Tay, Y., Rinn, J. & Pandolfi, P.P. The multilayered complexity of ceRNA crosstalk and competition. *Nature* **505**, 344–352 (2014).
27. Thomson, D.W. & Dinger, M.E. Endogenous microRNA sponges: Evidence and controversy. *Nat. Rev. Genet.* **17**, 272–283 (2016).
28. Yuan, J.H. *et al.* A long noncoding RNA activated by TGF- β promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell* **25**, 666–681 (2014).
29. Richards, E.J. *et al.* Long non-coding RNAs (LncRNA) regulated by transforming growth factor (TGF) β : LncRNA-hit-mediated TGF β -induced epithelial to mesenchymal transition in mammary epithelia. *J. Biol. Chem.* **290**, 6857–6867 (2015).
30. Xu, L. *et al.* Long non-coding RNA SMASR inhibits the EMT by negatively regulating TGF- β /SMAD signaling pathway in lung cancer. *Oncogene* **40**, 3578–3592 (2021).
31. Sakai, S. *et al.* Long noncoding RNA ELIT-1 acts as a SMAD3 cofactor to facilitate TGF- β /SMAD signaling and promote epithelial-mesenchymal transition. *Cancer Res.* **79**, 2821–2838 (2019).
32. Papoutsoglou, P. *et al.* The TGFB2-AS1 lncRNA regulates TGF- β signaling by modulating corepressor activity. *Cell Rep.* **28**, 3182–3198.e11 (2019).
33. Papoutsoglou, P. & Moustakas, A. Long non-coding RNAs and TGF- β signaling in cancer. *Cancer Sci.* **111**, 2672–2681 (2020).
34. Wang, P. *et al.* Long noncoding RNA lnc-TSI inhibits renal fibrogenesis by negatively regulating the TGF- β /SMAD3 pathway. *Sci. Transl. Med.* **10**, eaat2039 (2018).
35. Fan, C.N. *et al.* LncRNA LITATS1 suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation. *EMBO J.* **42**, e112806 (2023).
36. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
37. Wang, L. *et al.* CPAT: Coding-potential assessment tool using an alignment-free logistic regression model. *Nucleic Acids Res.* **41**, e74 (2013).
38. Dennler, S. *et al.* Direct binding of SMAD3 and SMAD4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. *EMBO J.* **17**, 3091–3100 (1998).
39. Marvin, D.L. *et al.* Dynamic visualization of TGF- β /SMAD3 transcriptional responses in single living cells. *Cancers (Basel)* **14**, 2508 (2022).
40. Padua, D. *et al.* TGF β primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**, 66–77 (2008).
41. Gil, N. & Ulitsky, I. Regulation of gene expression by cis-acting long non-coding RNAs. *Nat. Rev. Genet.* **21**, 102–117 (2020).
42. Sun, Q., Hao, Q. & Prasanth, K.V. Nuclear long noncoding RNAs: Key regulators of gene expression. *Trends Genet.* **34**, 142–157 (2018).
43. Yi, W. *et al.* CRISPR-assisted detection of RNA-protein interactions in living cells. *Nat. Methods* **17**, 685–688 (2020).
44. Zhu, L., Liu, Y., Tang, H. & Wang, P. FOXP3 activated-LINC01232 accelerates the stemness of non-small cell lung carcinoma by activating TGF- β signaling pathway and recruiting IGF2BP2 to stabilize TGFBR1. *Exp. Cell Res.* **413**, 113024 (2022).
45. Cheng, D.M. *et al.* Long noncoding RNA-SNHG20 promotes silica-induced pulmonary fibrosis by miR-490-3p/TGFBR1 axis. *Toxicology* **451**, 152683 (2021).
46. Hu, H.Y. *et al.* Long non-coding RNA TCONS_00814106 regulates porcine granulosa cell proliferation and apoptosis by sponging miR-1343. *Mol. Cell. Endocrinol.* **520**, 111064 (2021).
47. Li, Y.C., Zhao, Z., Sun, D. & Li, Y.F. Novel long noncoding RNA LINC02323 promotes cell growth and migration of ovarian cancer via TGF- β receptor 1 by miR-1343-3p. *J. Clin. Lab. Anal.* **35**, e23651 (2021).
48. Zhou, B., Guo, W.D., Sun, C.D., Zhang, B.Y. & Zheng, F. Linc00462 promotes pancreatic cancer invasiveness through the miR-665/TGFBR1-TGFBR2/SMAD2/3 pathway. *Cell Death Dis.* **9**, 706 (2018).
49. Jin, J., Jia, Z.H., Luo, X.H. & Zhai, H.F. Long non-coding RNA HOXA11-AS accelerates the progression of keloid formation via miR-124-3p/TGF β R1 axis. *Cell Cycle* **19**, 218–232 (2020).
50. Yang, G. & Lin, C.S. Long noncoding RNA SOX2-OT exacerbates hypoxia-induced cardiomyocytes injury by regulating miR-27a-3p/TGF β R1 axis. *Cardiovasc. Ther.* **2020**, 2016259 (2020).
51. Li, Y. *et al.* Long non-coding RNA SBF2-AS1 promotes hepatocellular carcinoma progression through regulation of miR-140-5p-TGFBR1 pathway. *Biochem. Biophys. Res. Commun.* **503**, 2826–2832 (2018).
52. Qi, J., Wu, Y.Y., Zhang, H.J. & Liu, Y.F. LncRNA NORAD regulates scar hypertrophy via miRNA-26a

- mediating the regulation of TGF β R1/2. *Adv. Clin. Exp. Med.* **30**, 395–403 (2021).
53. Gilbert, L.A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
 54. Ransohoff, J.D., Wei, Y.N. & Khavari, P.A. The functions and unique features of long intergenic non-coding RNA. *Nat. Rev. Mol. Cell Biol.* **19**, 143–157 (2018).
 55. Ulitsky, I. & Bartel, D.P. lincRNAs: Genomics, evolution, and mechanisms. *Cell* **154**, 26–46 (2013).
 56. Ogunjimi, A.A. *et al.* Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain. *Mol. Cell* **19**, 297–308 (2005).
 57. Kuratomi, G. *et al.* NEDD4-2 (neural precursor cell expressed, developmentally down-regulated 4-2) negatively regulates TGF- β (Transforming growth factor- β) signalling by inducing ubiquitin-mediated degradation of SMAD2 and TGF- β type I receptor. *Biochem. J.* **386**, 461–470 (2005).
 58. Hedrick, E., Mohankumar, K. & Safe, S. TGF β -induced lung cancer cell migration is NR4A1-dependent. *Mol. Cancer Res.* **16**, 1991–2002 (2018).
 59. Hedrick, E. & Safe, S. Transforming growth factor β /NR4A1-inducible breast cancer cell migration and epithelial-to-mesenchymal transition is p38 α (mitogen-activated protein kinase 14) dependent. *Mol. Cell. Biol.* **37**, e00306–e00317 (2017).
 60. Gao, Y. *et al.* LncRNA lncLy6C induced by microbiota metabolite butyrate promotes differentiation of Ly6C^{high} to Ly6C^{int/neg} macrophages through lncLy6C/C/EBP β /Nr4A1 axis. *Cell Discov.* **6**, 87 (2020).
 61. Stroud, J.C., Lopez-Rodriguez, C., Rao, A. & Chen, L. Structure of a TonEBP-DNA complex reveals DNA encircled by a transcription factor. *Nat. Struct. Biol.* **9**, 90–94 (2002).
 62. Lopez, A.M., Pegram, M.D., Slamon, D.J. & Landaw, E.M. A model-based approach for assessing in vivo combination therapy interactions. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13023–13028 (1999).
 63. Lee, S.D., Woo, S.K. & Kwon, H.M. Dimerization is required for phosphorylation and DNA binding of TonEBP/NFAT5. *Biochem. Biophys. Res. Commun.* **294**, 968–975 (2002).
 64. Guo, K. & Jin, F.G. NFAT5 promotes proliferation and migration of lung adenocarcinoma cells in part through regulating AQP5 expression. *Biochem. Biophys. Res. Commun.* **465**, 644–649 (2015).
 65. Li, J.T. *et al.* Nuclear factor of activated T cells 5 maintained by Hotair suppression of miR-568 upregulates S100 calcium binding protein A4 to promote breast cancer metastasis. *Breast Cancer Res.* **16**, 454 (2014).
 66. Meng, X., Li, Z., Zhou, S., Xiao, S. & Yu, P. miR-194 suppresses high glucose-induced non-small cell lung cancer cell progression by targeting NFAT5. *Thorac. Cancer* **10**, 1051–1059 (2019).
 67. Yang, M.J., Ke, H.G. & Zhou, W. LncRNA RMRP promotes cell proliferation and invasion through miR-613/NFAT5 axis in non-small cell lung cancer. *Oncotargets Ther.* **13**, 8941–8950 (2020).
 68. Lambert, S.A. *et al.* The human transcription factors. *Cell* **172**, 650–665 (2018).
 69. Liu, J. *et al.* TF-PROTACs enable targeted degradation of transcription factors. *J. Am. Chem. Soc.* **143**, 8902–8910 (2021).
 70. Shao, J.W. *et al.* Destruction of DNA-binding proteins by programmable oligonucleotide PROTAC (O^{PROTAC}): Effective targeting of LEF1 and ERG. *Adv. Sci. (Weinh)* **8**, 2102555 (2021).
 71. Maruyama, R. & Yokota, T. Knocking down long noncoding RNAs using antisense oligonucleotide gapmers. *Methods Mol. Biol.* **2176**, 49–56 (2020).
 72. Dey, S.K. & Jaffrey, S.R. RIBOTACs: Small molecules target RNA for degradation. *Cell Chem. Biol.* **26**, 1047–1049 (2019).
 73. Fan, C. *et al.* OVOL1 inhibits breast cancer cell invasion by enhancing the degradation of TGF- β type I receptor. *Signal Transduct. Target. Ther.* **7**, 126 (2022).
 74. Wang, Q. *et al.* Broadening the reach and investigating the potential of prime editors through fully viral gene-deleted adenoviral vector delivery. *Nucleic Acids Res.* **49**, 11986–12001 (2021).
 75. Sinha, A. *et al.* Visualizing dynamic changes during TGF- β -induced epithelial to mesenchymal transition. *Methods Mol. Biol.* **2488**, 47–65 (2022).
 76. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
 77. Daniel Salas-Lloret, C.v.d.M., Easa Nagamalleswari, Ekaterina Gracheva, Arnoud H. de Ru, H. Anne Marie Otte, Peter A. van Veelen, Andrea Pichler, Joachim Goedhart, Alfred C.O. Vertegaal, Román González-Prieto SUMO activated target traps (SATTs) enable the identification of a comprehensive E3-specific SUMO proteome. *bioRxiv* 2022.2006.2022.497173, (2022).
 78. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
 79. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
 80. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).

81. Ren, J., Liu, S., Cui, C. & Ten Dijke, P. Invasive behavior of human breast cancer cells in embryonic zebrafish. *J. Vis. Exp.* **122**, 55459 (2017).
82. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

Supplementary information

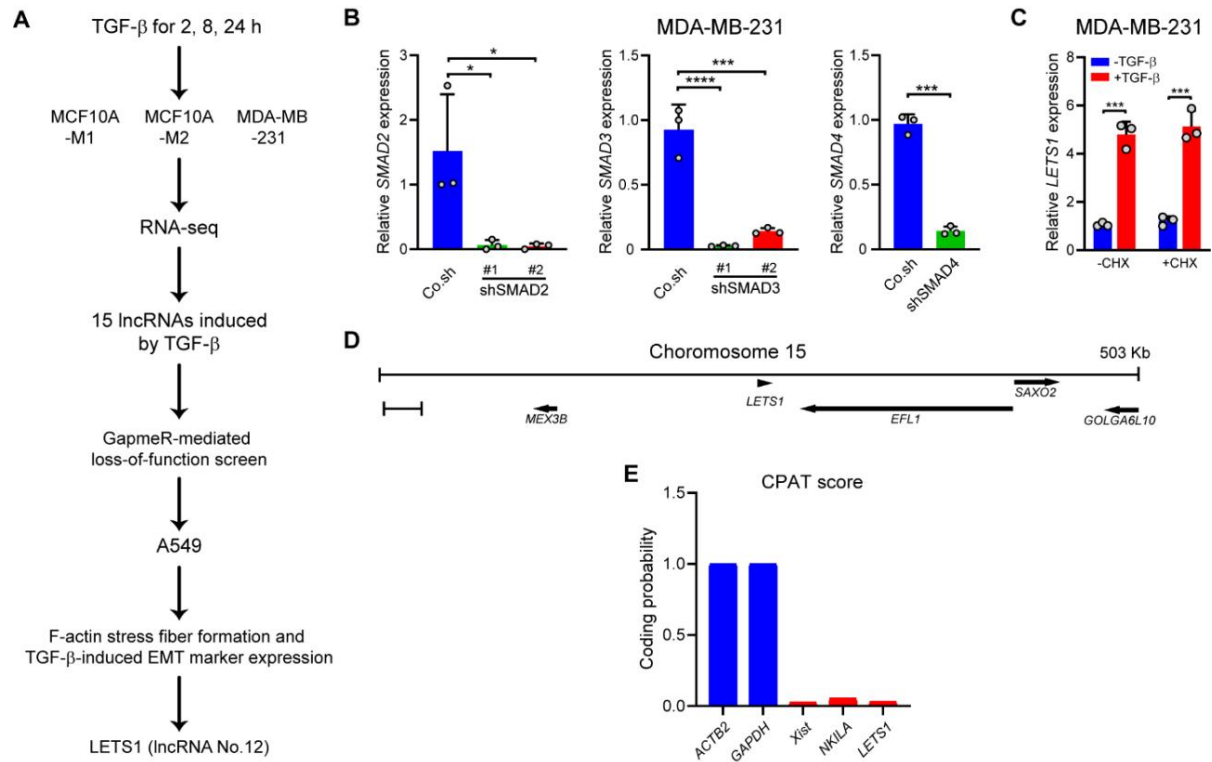


Fig. S1. *LETS1* is a TGF- β -induced lncRNA. (A) Workflow showing the screening strategy to identify TGF- β -induced *LETS1* that functions as an activator of EMT. (B) Expression of *SMAD2*, *SMAD3* or *SMAD4* in MDA-MB-231 cells upon shRNA-mediated *SMAD2*, *SMAD3*, or *SMAD4* knockdown. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (C) Quantification of *LETS1* expression in MDA-MB-231 cells pre-treated with cycloheximide and stimulated with TGF- β or vehicle. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of two independent experiments. (D) Schematic representation of the genomic location of *LETS1* and its neighboring genes. The arrows indicate the direction of transcription. (E) The predicted coding potential of protein-coding mRNAs (*ACTB2* and *GAPDH*), well-annotated lncRNAs (*Xist* and *NKILA*) and *LETS1*. In (B, left and middle), significance was assessed using one-way ANOVA followed by Dunnett's multiple comparisons test. In (B, right) and (C), significance was assessed using unpaired Student's t test. *, $0.01 < p < 0.05$; ***, $0.0001 < p < 0.001$; ****, $p < 0.0001$.

The lncRNA *LETS1* promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism

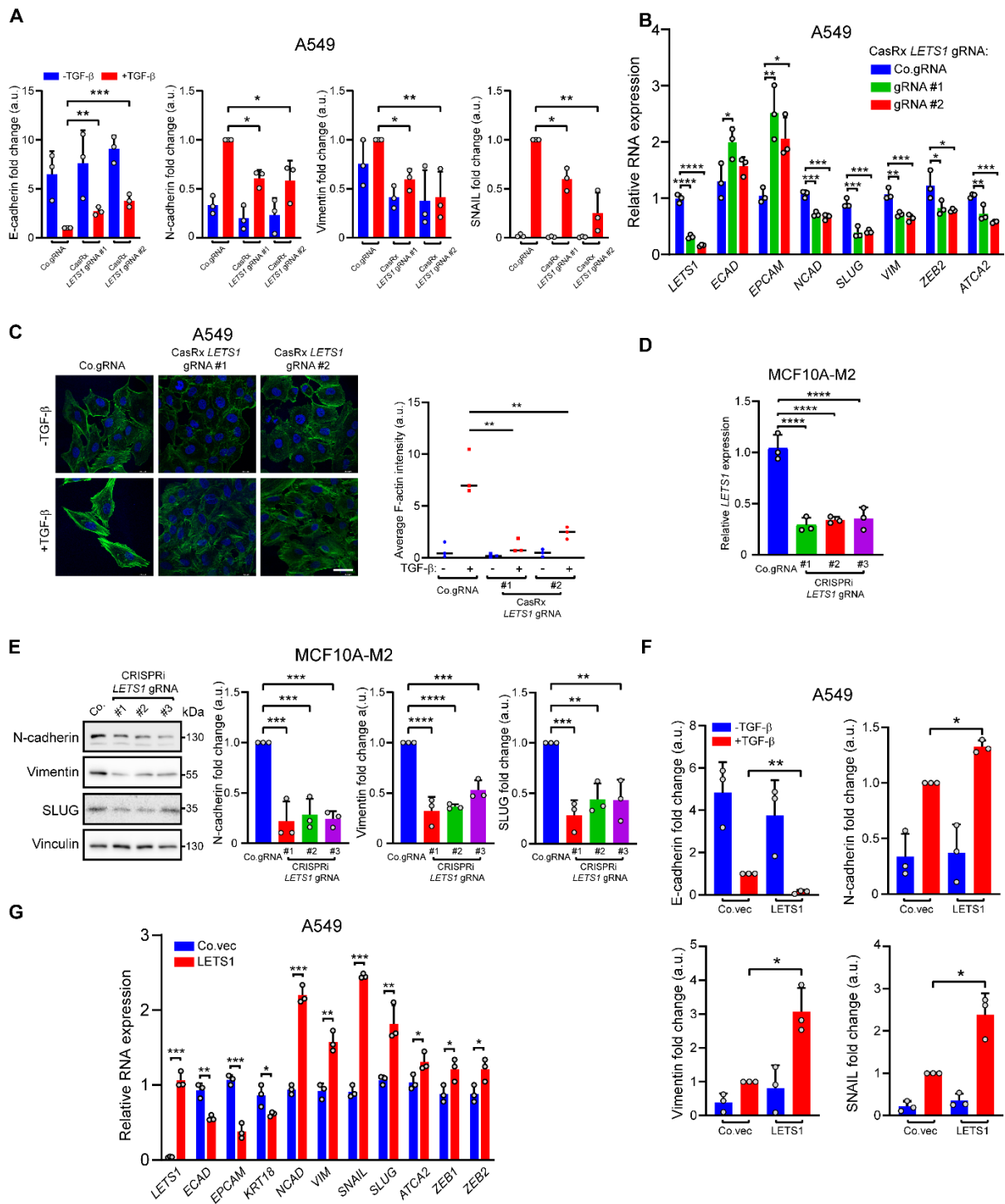


Fig. S2. *LETS1* promotes TGF- β -induced EMT. (A) Quantification results of western blotting in Fig. 2A. Statistical results are means \pm SD from three independent experiments. a.u., arbitrary units. (B) Expression of EMT markers in A549 cells upon CRISPR/CasRx-mediated *LETS1* knockdown. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (C) Fluorescent staining for F-actin in A549 cells upon *LETS1* depletion by CRISPR/CasRx. Co.gRNA, empty vector. DAPI staining was performed to visualize nuclei. Scale bar, 38.8 μ m. Quantification of average F-actin intensity is shown as means \pm SD from three independent experiments. (D) *LETS1* expression in MCF10A-M2 cells upon CRISPRi-mediated *LETS1* knockdown. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (E) Immunoblotting for N-cadherin, Vimentin, and SLUG in MCF10A-M2 cells expressing the CRISPRi construct and empty vector (Co.) or *LETS1*-targeting guide RNA (gRNA). Vinculin, loading control. Quantification results are shown as means \pm SD from three independent experiments. (F) Quantification results of western blotting in Fig. 2B. Statistical results are means \pm SD from three independent experiments. (G) Expression of EMT markers in A549 cells upon *LETS1* ectopic expression. RT-qPCR results are shown as means \pm SD from three biological replicates in one

independent experiment and representative of at least three independent experiments. In (A), (E) and (F), significance was assessed using paired Student's *t* test. In (B), (C) and (D), significance was assessed using one-way ANOVA followed by Dunnett's multiple comparisons test. In (G), significance was assessed using unpaired Student's *t* test. *, 0.01 < *p* < 0.05; **, 0.001 < *p* < 0.01; ***, 0.0001 < *p* < 0.001; ****, *p* < 0.0001.

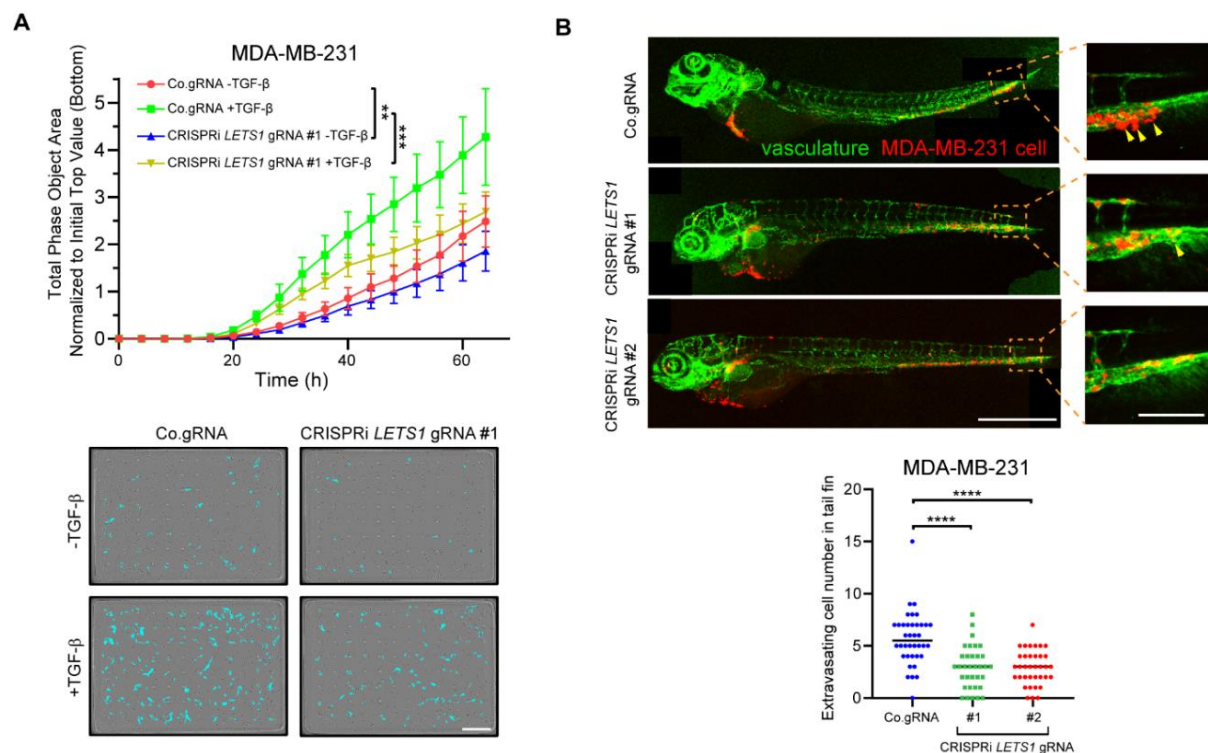


Fig. S3. *LETS1* knockdown attenuates TGF- β -induced cell migration and extravasation. (A) An IncuCyte chemotactic migration assay was performed with MDA-MB-231 cells upon CRISPRi-mediated *LETS1* depletion. Cells that migrated to the bottom of chambers are marked in blue in the images. The migration results are expressed as means \pm SD from six biological replicates in one independent experiment and representative of at least three independent experiments. Scale bar, 400 μ m. Significance was assessed using two-way ANOVA followed by Tukey's multiple comparisons test. **, 0.001 < *p* < 0.01; ***, 0.0001 < *p* < 0.001. (B) In vivo zebrafish extravasation experiments with MDA-MB-231 cells upon CRISPRi-mediated *LETS1* depletion. Extravasated breast cancer cells in the zoomed tail fin area are indicated with yellow arrows. Numbers of extravasated cell are expressed as means \pm SD. Scale bars, 309.1 μ m (whole fish); 154.5 μ m (enlargements). N = at least 30 fish per treatment group. Images are representative of two independent experiments. Significance was assessed using one-way ANOVA followed by Dunnett's multiple comparisons test. ****, *p* < 0.0001. *LETS1* knockdown efficiency is shown in fig. S5A.

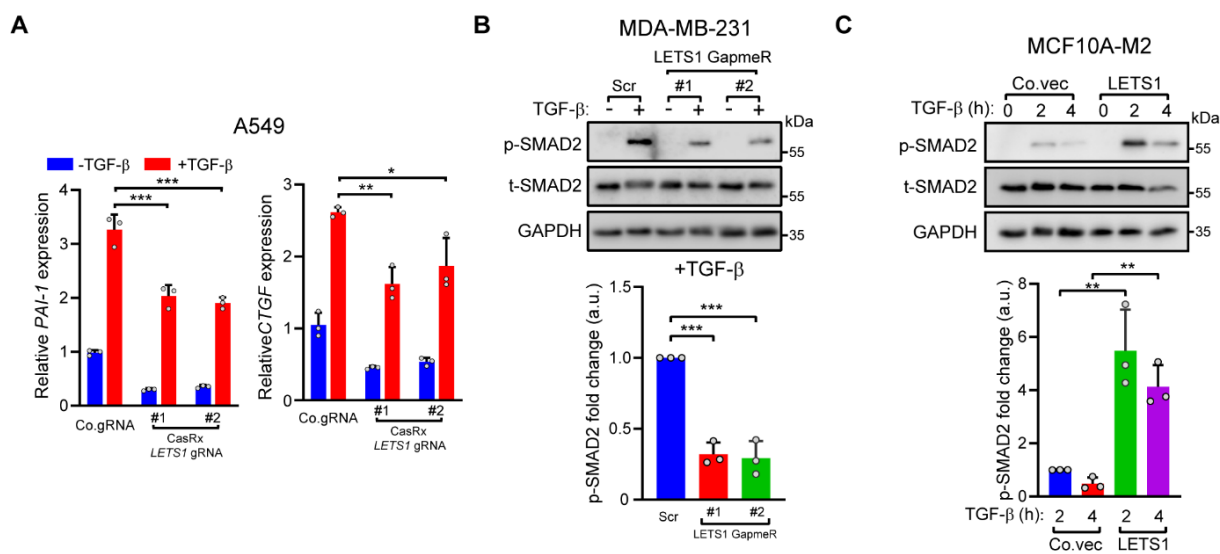


Fig. S4. *LETS1* knockdown attenuates TGF- β /SMAD signaling. (A) *PAI-1* and *CTGF* expression in A549 cells upon

The lncRNA *LETS1* promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism

CRISPR/CasRx-mediated *LETS1* depletion. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (B, C) Immunoblotting for phosphorylated (p-) and total (t-) SMAD2 in TGF- β -stimulated MDA-MB-231 or MCF10AM2 cells in which *LETS1* was knocked down by GapmeR (MDA-MB-231) or in which *LETS1* was overexpressed (MCF10A-M2) Blots are representative of at least three independent experiments and statistical results are means \pm SD from three independent experiments. In (A), significance was assessed using one-way ANOVA followed by Dunnett's multiple comparisons test. In (B) and (C), significance was assessed using paired Student's *t* test. *, 0.01 < *p* < 0.05; **, 0.001 < *p* < 0.01; ***, 0.0001 < *p* < 0.001.

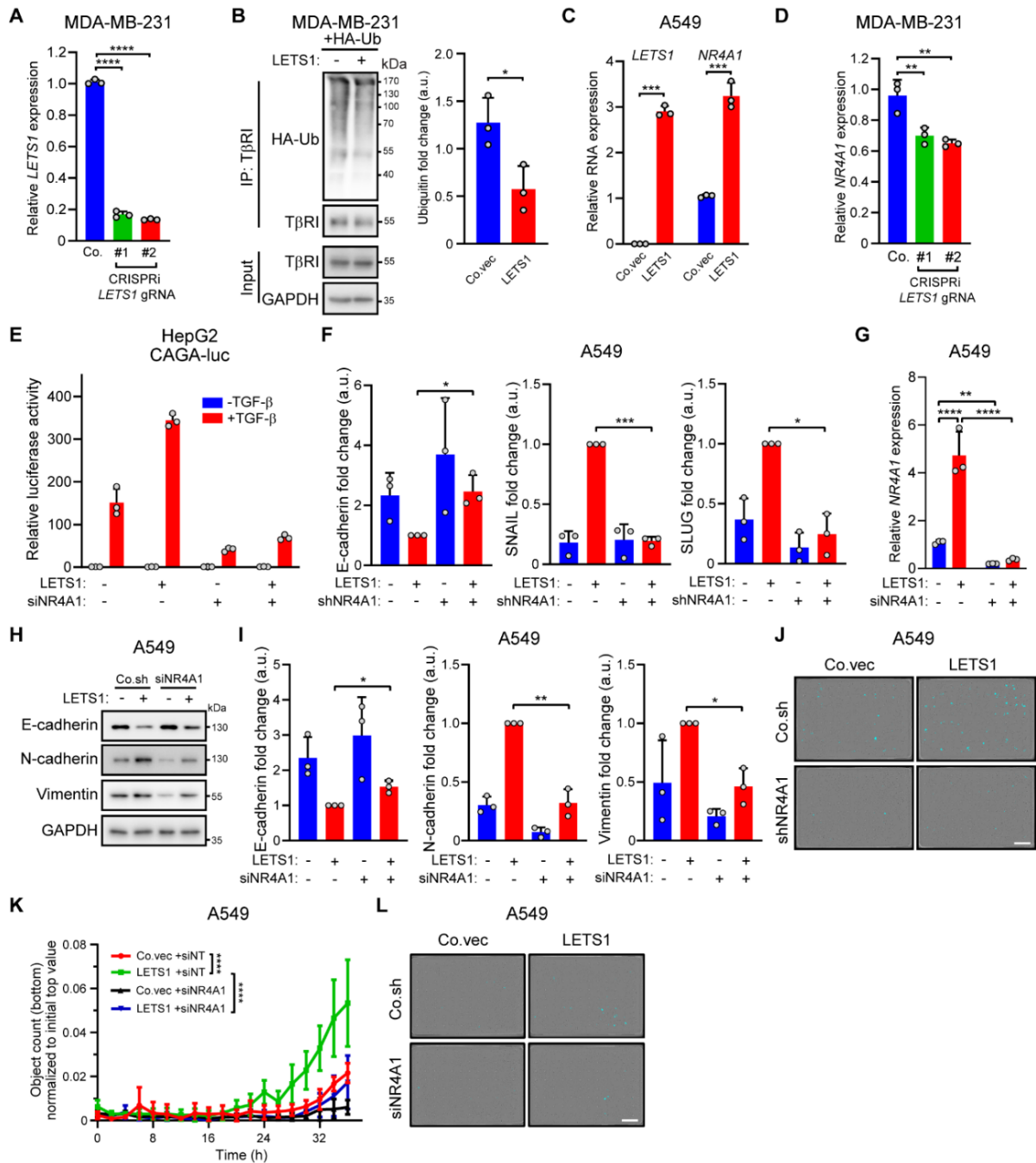


Fig. S5. *NR4A1* is induced by *LETS1*. (A) *LETS1* expression in MDA-MB-231 cells upon CRISPRi-mediated *LETS1* knockdown. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (B) Immunoblotting for HA and T β RI in total lysates (input) and T β RI immunoprecipitates (IP) from MDA-MB-231 cells expressing HA-ubiquitin (HA-Ub) and empty vector (Co.vec) or *LETS1*. Ubiquitin was quantified in the T β RI immunoprecipitates. Quantitative data are means \pm SD from three independent experiments. (C) *LETS1* and *NR4A1* expression in A549 cells upon ectopic *LETS1* expression. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. (D) *NR4A1* expression in MDA-MB-231 cells upon CRISPRi-mediated *LETS1* knockdown. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three

independent experiments. (E) Luciferase activity in TGF- β -stimulated HepG2 cells transfected with the expression construct for the SMAD3/4 transcriptional reporter CAGA-luc plus the LETS1 ectopic expression construct and the NR4A1-targeting siRNA as indicated. The relative luciferase activities are representative of at least three independent experiments and expressed as means \pm SD from three wells with cells per treatment group in one experiment. (F) Quantification results of western blotting in Fig. 4G. Statistical results are means \pm SD from three independent experiments. (G) *NR4A1* expression in MDA-MB-231 cells with LETS1 ectopic expression and siRNA-mediated NR4A1 knockdown. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of two independent experiments. (H) Immunoblotting for E-cadherin, N-cadherin, and Vimentin in A549 cells in which LETS1 was overexpressed and NR4A1 was knocked down by siRNA as indicated. Blots are representative of three independent experiments. (I) Quantification results of western blotting in fig. S5H. Statistical results are means \pm SD from three independent experiments. (J) The representative images of migrated cells in Fig. 4H. The cells that migrated to the bottom of chambers are marked in blue in the images. Scale bar, 400 μ m. (K) Quantification of migrated cells in IncuCyte chemotactic migration assays using A549 cells with LETS1 overexpression and NR4A1 knockdown by siRNA as indicated. The migration results are expressed as means \pm SD from five biological replicates in one independent experiment and representative of two independent experiments. (L) The representative images of migrated cells in fig. S5K. The cells that migrated to the bottom of chambers are marked in blue in the images. Scale bar, 400 μ m. In (A), (D), and (G), significance was assessed using one-way ANOVA followed by Dunnett's multiple comparisons test. In (C), significance was assessed using unpaired Student's *t* test. In (B), (F) and (I), significance was assessed using paired Student's *t* test. In (K), significance was assessed using two-way ANOVA followed by Tukey's multiple comparisons test. *, 0.01 < *p* < 0.05; **, 0.001 < *p* < 0.01; ***, 0.0001 < *p* < 0.001; ****, *p* < 0.0001.

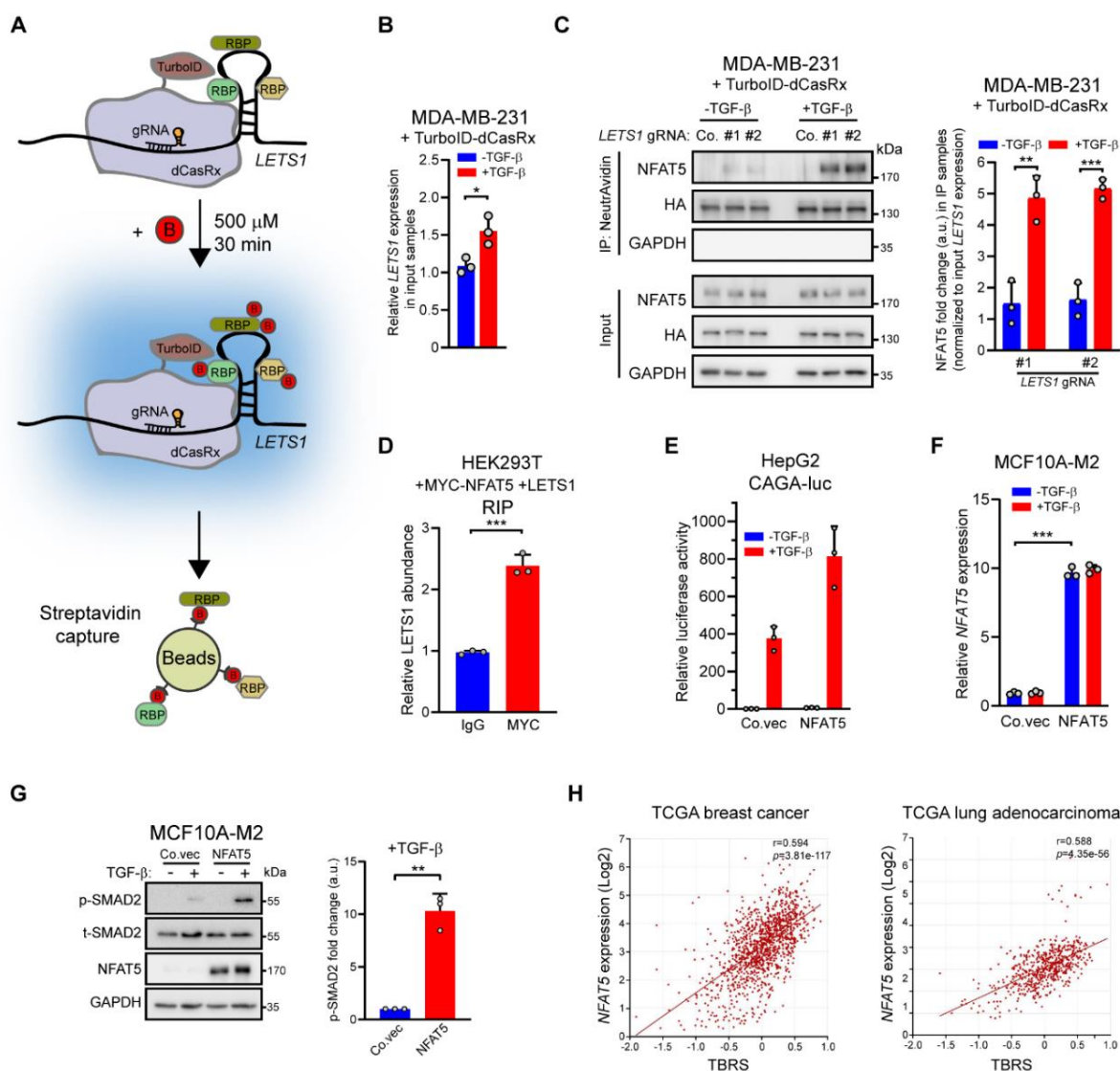


Fig. S6. NFAT5 potentiates TGF- β -SMAD signaling. (A) Scheme of the CARPID workflow. B: biotin; RBP: RNA binding protein. (B) *LETS1* expression in MDA-MB-231 cells with TurboID-dCasRx expression and stimulated with TGF- β or vehicle. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of two independent experiments. (C) Interactions between LETS1 and NFAT5 were analyzed by CARPID in MDA-MB-231 with TurboID-dCasRx expression. Western blotting was performed to detect NFAT5 expression in whole-cell lysates (Input)

The lncRNA *LETS1* promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism

and immunoprecipitates (IP). Blots are representative of at least three independent experiments. NFAT5 abundance in IP samples was normalized to *LETS1* expression in input samples. Statistical results are means \pm SD from three independent experiments. **(D)** RIP assay quantifying *LETS1* abundance in MYC immunoprecipitates from HEK293T cells with MYCN4A1 and *LETS1* ectopic expression. *LETS1* abundance in MYC immunoprecipitates is presented as relative to that in IgG immunoprecipitates. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of two independent experiments **(E)** Luciferase activity in TGF- β -stimulated HepG2 cells transfected with the expression construct for the SMAD3/4 transcriptional reporter CAGA-luc plus the NFAT5 ectopic expression construct. The relative luciferase activities are representative of at least three independent experiments and expressed as means \pm SD from three wells with cells per treatment group in one experiment. **(F)** *NFAT5* expression in MCF10A-M2 cells with NFAT5 ectopic expression. The RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of at least three independent experiments. **(G)** Immunoblotting for phosphorylated (p-) and total (t-) SMAD2 and NFAT5 in TGF- β -stimulated MCF10A-M2 cells with NFAT5 overexpression. Quantitative data show the abundance of p-SMAD2 relative to t-SMAD2. Results are means \pm SD from three independent experiments. **(H)** Correlations between NFAT5 and the TGF- β gene response signature (TBRS) in patients with breast cancer or lung adenocarcinoma. In (B), (D), and (F), significance was assessed using unpaired Student's *t* test. In (C) and (G), significance was assessed using paired Student's *t* test. In (H), statistical analysis was performed using Pearson's correlation test. *, 0.01 < *p* < 0.05; **, 0.001 < *p* < 0.01; ***, 0.0001 < *p* < 0.001.

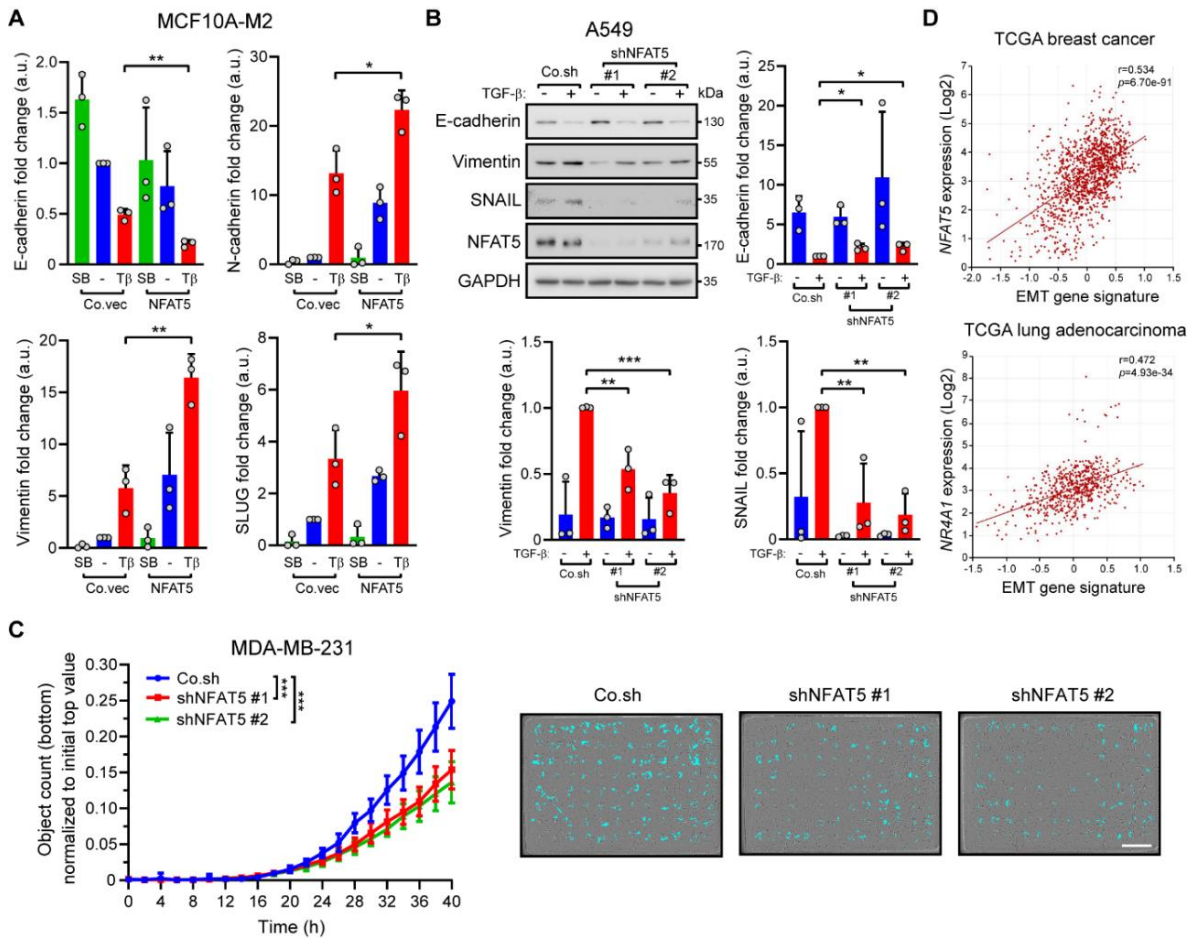


Fig. S7. NFAT5 potentiates TGF- β -induced EMT and cell migration. **(A)** Quantification results of western blotting in Fig. 5H. Statistical results are means \pm SD from three independent experiments. **(B)** Immunoblotting for E-cadherin, Vimentin, SNAIL, and NFAT5 in A549 cells upon NFAT5 knockdown and treated with vehicle or TGF- β (T β). Co.sh, empty vector. Blots are representative of at least three independent experiments. Quantification results are means \pm SD from three independent experiments. **(C)** Quantification of migrated cells in IncuCyte chemotactic migration assays using MDA-MB231 cells upon NFAT5 knockdown. The cells that migrated to the bottom of chambers are marked in blue in the images. The migration results are expressed as means \pm SD from 12 biological replicates in one independent experiment and representative of two independent experiments. Scale bar, 400 μ m. **(D)** Correlations between *NFAT5* and the EMT signature in patients with breast cancer or lung adenocarcinoma. In (A) and (B), significance was assessed using paired Student's *t* test. In (C), significance was assessed using two-way ANOVA followed by Tukey's multiple comparisons test. In (D), statistical analysis was performed using Pearson's correlation test. *, 0.01 < *p* < 0.05; **, 0.001 < *p* < 0.01; ***, 0.0001 < *p* < 0.001.

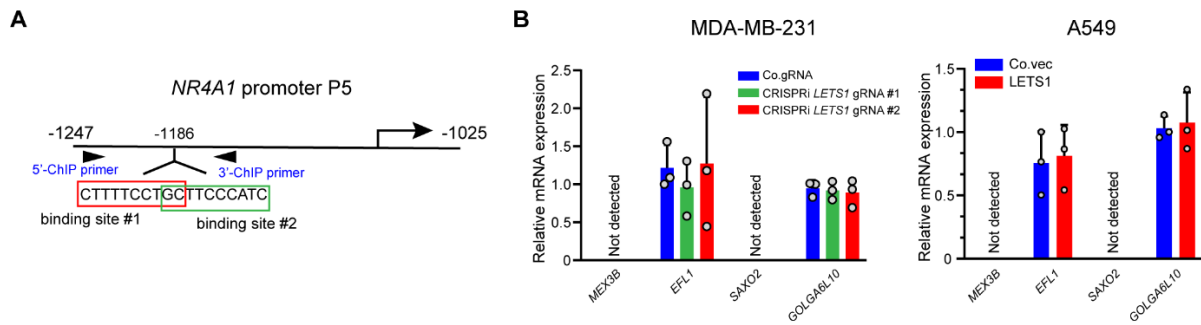


Fig. S8. The expression of *LETS1* nearby genes is not affected upon *LETS1* misexpression. (A) Schematic representation of *NR4A1* promoter P5. The two putative binding sites of NFAT5 and the binding sites of ChIP primers are shown. (B) Quantification of *LETS1* neighboring gene expression in A549 cells with *LETS1* ectopic expression and MDA-MB-231 cells upon CRISPRi-mediated *LETS1* knockdown. RT-qPCR results are shown as means \pm SD from three biological replicates in one independent experiment and representative of two independent experiments.

Supplementary tables are online at <https://www.science.org/doi/10.1126/scisignal.adf1947>.



Chapter 4

OVOL1 inhibits breast cancer cell invasion by enhancing the degradation of TGF- β type I receptor

Chuannan Fan^{1,2}, Qian Wang¹, Gerard van der Zon^{1,2}, Jiang Ren¹, Cedrick Agaser³, Roderick C Sliker^{1,4}, Prasanna Vasudevan Iyengar^{1,2}, Hailiang Mei³, and Peter ten Dijke^{1,2*}

¹Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands.

²Oncode Institute, Leiden University Medical Center, Leiden, The Netherlands.

³Department of Biomedical Data Sciences, Sequencing Analysis Support Core, Leiden University Medical Center, Leiden, The Netherlands.

⁴Department of Epidemiology and Data Science, Amsterdam Public Health Institute, Amsterdam Cardiovascular Sciences Institute, Amsterdam UMC, location VUmc, Amsterdam, the Netherlands.

*Correspondence: Peter ten Dijke

Department of Cell and Chemical Biology and Oncode Institute, Leiden University Medical Center, Einthovenweg 20, 2300 RC, Leiden, The Netherlands

Email: p.ten_dijke@lumc.nl; Telephone: +31 71526 9271; Fax: +31 71 526 8270

Abstract

Ovo like transcriptional repressor 1 (OVOL1) is a key mediator of epithelial lineage determination and mesenchymal-epithelial transition (MET). The cytokines transforming growth factor- β (TGF- β) and bone morphogenetic proteins (BMP) control the epithelial-mesenchymal plasticity (EMP) of cancer cells, but whether this occurs through interplay with OVOL1 is not known. Here we show that OVOL1 is inversely correlated with the epithelial-mesenchymal transition (EMT) signature, and is an indicator of a favorable prognosis for breast cancer patients. OVOL1 suppresses EMT, migration, extravasation and early metastatic events of breast cancer cells. Importantly, BMP strongly promotes the expression of OVOL1, which enhances BMP signaling in turn. This positive feedback loop is established through the inhibition of TGF- β receptor signaling by OVOL1. Mechanistically, OVOL1 interacts with and prevents the ubiquitination and degradation of SMAD family member 7 (SMAD7), which is a negative regulator of TGF- β type I receptor stability. Moreover, a small molecule compound 6-formylindolo(3,2-b)carbazole (FICZ) was identified to activate OVOL1 expression and thereby antagonizes (at least in part) TGF- β -mediated EMT and migration in breast cancer cells. Our results uncover a novel mechanism by which OVOL1 attenuates TGF- β /SMAD signaling and maintains the epithelial identity of breast cancer cells.

Introduction

Breast cancer is one of the most commonly diagnosed cancers among females worldwide¹. Although recent developed therapies have improved the treatment of breast cancer, the survival rate decreases dramatically if patients develop distant metastases². The epithelial-mesenchymal transition (EMT) process plays a critical role in the invasion-metastasis cascade³. During this process, normal epithelial cells lose cell-cell junctions and acquire the fibroblast-like morphology, enhanced migratory and invasive properties^{4, 5}. In addition, mesenchymal cancer cells are less sensitive to chemotherapy^{6, 7}. At the molecular level, the classic epithelial marker E-cadherin (*ECAD*), which contributes to the cellular adhesion, is transcriptionally repressed by core EMT transcription factors including SNAIL, ZEB and TWIST family members⁸⁻¹⁰. Moreover, the expression of mesenchymal markers such as Fibronectin (*FN*), N-cadherin (*NCAD*), Vimentin (*VIM*) and α -SMA (*ACTA2*) is upregulated, leading to the cytoskeleton reconstruction and cell migration³. Of note, during the highly dynamic EMT process, a large proportion of cells, which share both epithelial and mesenchymal characteristics, stay in an intermediate and reversible state. The acquisition of this hybrid state is described as epithelial-mesenchymal plasticity (EMP), which is also referred to as partial EMT^{11, 12}.

The cytokine transforming growth factor- β (TGF- β) is one of the most crucial drivers to induce EMT⁵. TGF- β initiates cellular responses by specific binding to cell surface TGF- β type II receptor (T β RII) and TGF- β type I receptor (T β RI). Upon T β RII mediated phosphorylation of T β RI, the latter receptor is activated and induces the phosphorylation and activation of regulatory (R)-SMADs, i.e., SMAD2 and SMAD3. Phosphorylated R-SMADs, upon forming a complex with SMAD4, translocate into the nucleus to activate the expression of typical target genes such as plasminogen activator inhibitor-type 1 (*PAI-1*) and connective tissue growth factor (*CTGF*)¹³⁻¹⁵. Notably, core EMT transcription factors like *SNAIL*, *SLUG* and *ZEB1/2* can be directly induced by the SMAD complex and contribute to TGF- β -mediated changes of aforementioned EMT markers and the reorganization of cytoskeleton such as filamentous (F)-actin⁵.

Bone morphogenetic proteins (BMPs), which are cytokines belonging to the TGF- β family, activate the signaling by inducing complex formation of BMP type II receptor (BMPRII) and

BMP type I receptor (BMPRI). Afterwards, BMPRI recruits and phosphorylates SMAD1/5/8, which translocate into the nucleus together with SMAD4 and promote the transcription of target genes such as inhibitor of DNA binding 1 (*ID1*) and inhibitor of DNA binding 3 (*ID3*)¹⁶. BMP signaling has been unveiled to maintain epithelial identity and attenuate the metastatic potential of breast cancer cells, which can be achieved by counteracting TGF- β signaling¹⁷. Moreover, TGF- β is able to antagonize BMP signaling, indicating the imbalance between TGF- β and BMP may influence the EMT status and invasive abilities of cancer cells¹⁸⁻²⁰.

To finely control the propagation of TGF- β signaling, various negative feedback loops and multiple layers of regulation exist²¹. One of these loops is accomplished by the transcriptional activation of inhibitory (I)-SMAD7. SMAD7 protein is primarily localized in the nucleus, where it interacts with the SMAD specific E3 ubiquitin protein ligases (SMURFs)^{22, 23}. In response to TGF- β , the SMAD7/SMURFs complex translocate into the cytoplasm and interacts with T β RI, leading to the proteasome-mediated degradation of T β RI²²⁻²⁴. Furthermore, SMAD7 itself is polyubiquitinated and degraded by several E3 ligases including ARKADIA and RNF12^{25, 26}, whereas some deubiquitylating enzymes (DUBs) including USP26 and OTUD1 are capable of removing the polyubiquitin chains off SMAD7^{27, 28}.

Ovo like (OVOL) proteins, among which OVOL1 and OVOL2 are better investigated than OVOL3, are pivotal determinants of epithelial lineage and differentiation during embryonic development²⁹⁻³². Structurally, both OVOL1 and OVOL2 consist of an N-terminal SNAIL1/GFI (SNAG) domain, which acts as a molecular hook to recruit protein partners such as histone deacetylases (HDACs), and four zinc finger domains, which are responsible for the DNA binding capacity^{32, 33}. Due to the structural similarity and identical DNA binding sequence, OVOL1 and OVOL2 have been reported to function redundantly, e.g. in regulating epithelial plasticity and differentiation of epidermal progenitor cells³⁰. In addition, OVOL1 can suppress the transcription of *OVOL2* and itself^{33, 34}. In mesenchymal breast and prostate cancer cells, ectopic expression of OVOL1 or/and OVOL2 can induce mesenchymal-epithelial transition (MET), which is the reverse process of EMT, thereby suppressing cell migration³⁵. Although OVOL2 has been uncovered to attenuate TGF- β -induced EMT in breast cancer cells³⁶, the function of OVOL1 in the regulation of TGF- β /BMP signal transduction during breast cancer progression is ill-defined. Here, we identify *OVOL1* as a potent downstream target of BMP/SMAD and lesser extent of TGF- β /SMAD signaling, whose activities are in turn regulated by OVOL1 in a positive and negative feedback manner, respectively. Furthermore, we elucidate the mechanism by which OVOL1 attenuates TGF- β signaling and breast cancer metastasis. Importantly, OVOL1 interacts with and stabilizes SMAD7. Moreover, 6-formylindolo(3,2-b)carbazole (FICZ) was identified as a compound to potently activate OVOL1 expression, which may offer therapeutic potential for breast cancer patients.

Results

***OVOL1* is inversely correlated with EMT and is associated with favorable clinical outcomes in breast cancer patients**

Since OVOL transcriptional repressors have been reported to potentiate MET in prostate and breast cancer cells³⁵, we asked whether the expression of three *OVOL* genes was dysregulated during breast cancer progression. A dataset from 51 breast cancer cell lines revealed that *OVOL1* mRNA levels were downregulated in cell lines classified into an aggressive basal subtype, compared with those cell lines grouped into a less aggressive luminal subtype (Supplementary Fig. 1a)³⁷. However, no difference of *OVOL2* and *OVOL3* expression between these two subgroups was observed (Supplementary Fig. 1a). This result hints that *OVOL1*, in comparison with the other two *OVOL* members, may play an unique role in breast cancer

progression.

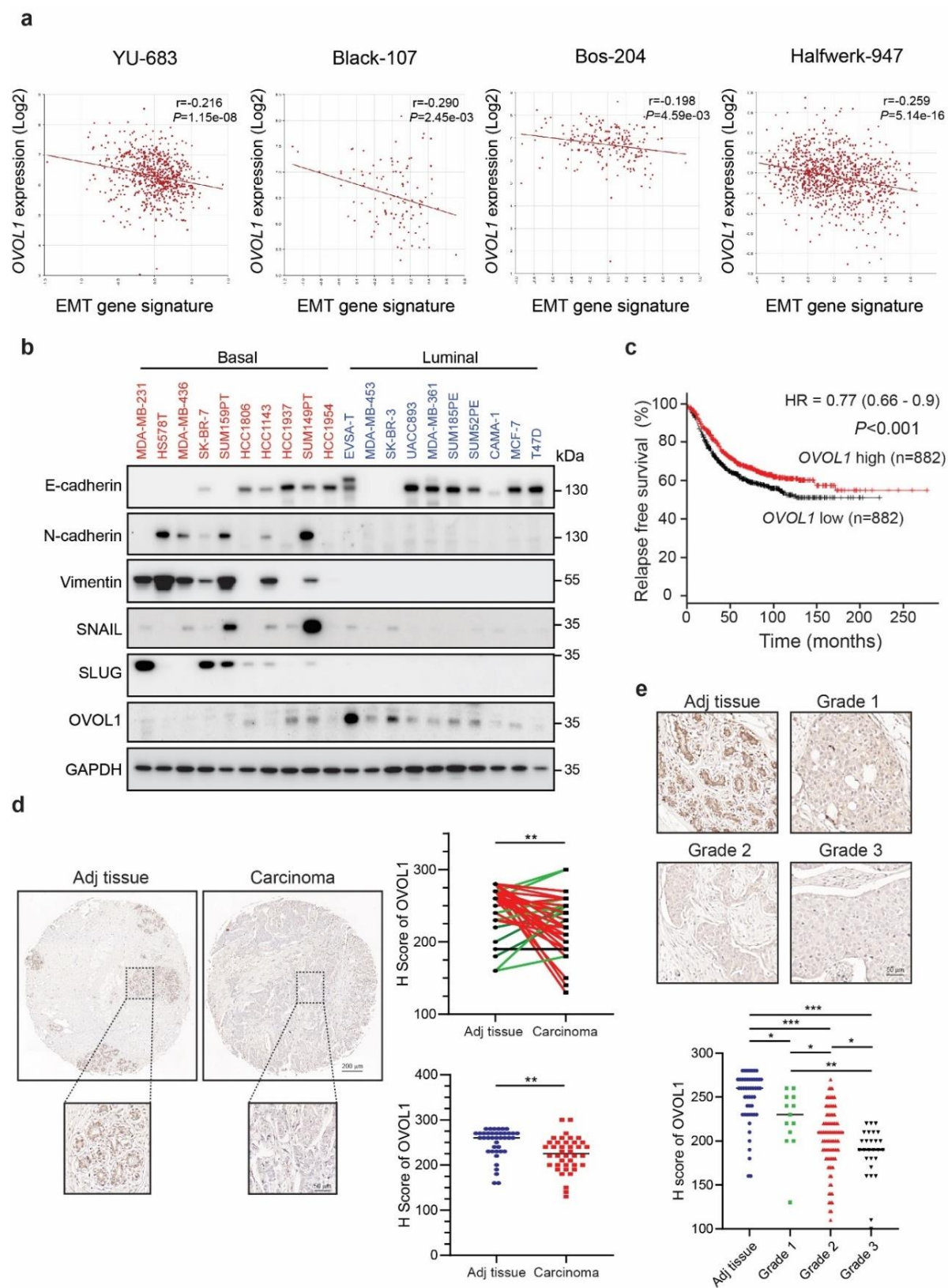


Fig. 1 OVOL1 expression is inversely correlated with the EMT gene signature and associated with a favorable prognosis in breast cancer patients. (a) Scatter plot displaying inverse correlation between *OVOL1* expression and the EMT gene signature in four breast cancer datasets. Titles on top of each panels indicate the datasets in which the RNA-seq data were analyzed. Pearson's correlation coefficient tests were performed to assess the statistical significance. (b) Western blotting

detection of EMT markers and OVOL1 levels in 10 basal and 10 luminal breast cancer cell lines. GAPDH levels were analyzed to control for equal loading. (c) Kaplan-Meier survival curves illustrating the relapse free survival of breast cancer patients stratified by *OVOL1* expression. Data were generated from Kaplan-Meier Plotter (<https://kmpplot.com/analysis/>). (d) Representative images of OVOL1 immunohistochemistry results in breast invasive ductal carcinoma (Carcinoma; n=40) and matched cancer adjacent breast tissues (Adj tissue; n=40) (upper). Comparisons between the H score of OVOL1 in the paired tissues (right upper) or unpaired Adj versus Carcinoma tissues (right lower) were performed. Paired tissues with higher OVOL1 expression in Adj tissues compared with Carcinoma are highlighted as red, whereas paired tissues with lower OVOL1 expression in Adj tissues compared with Carcinoma are highlighted as green. The results are expressed as mean \pm SD. ** 0.001 < p < 0.01. (e) Representative images of OVOL1 immunohistochemistry staining results in breast invasive carcinoma tissues with different grades (Grade 1, n=13; Grade 2, n=95; Grade 3, n=27) and cancer adjacent breast tissues (Adj tissue; n=50) (upper). Comparison results of OVOL1 H scores in Adj tissues and different groups of carcinoma tissues are shown in the lower panel. The results are expressed as mean \pm SD. * 0.01 < p < 0.05, ** 0.001 < p < 0.01, *** 0.0001 < p < 0.001.

Since it is pivotal for epithelial cells to acquire the EMP ability to initiate the invasion-metastasis cascade⁴, we examined the possible association between *OVOL1* expression and an established set of EMT genes³⁸ in breast cancer patients. Breast cancer patients with a higher EMT score were considered as more mesenchymal-like, while those with a lower EMT score were considered as more epithelial-like. Interestingly, a significant inverse correlation between *OVOL1* mRNA expression and the EMT signature was observed in four breast cancer patient cohorts (Fig. 1a). Moreover, we found positive correlations between *OVOL1* and epithelial markers (*ECAD*, *EPCAM* and *KRT18*), and negative correlations between *OVOL1* and mesenchymal markers (*NCAD*, *VIM* and *FN*), in a breast cancer patient cohort³⁹ and the aforementioned 51 cell line dataset³⁷ (Supplementary Fig. 1b, c). E-cadherin and OVOL1 were more prominently expressed, while the mesenchymal markers, including N-cadherin, Vimentin, SNAIL and SLUG, were lower expressed in 10 luminal cell lines than 10 basal cell lines (Fig. 1b). In addition, *OVOL1* mRNA and protein levels were determined in commonly used breast cell lines with either epithelial and/or mesenchymal features, including normal breast epithelial cells MCF10A-M1, premalignant breast cells MCF10A-M2 and luminal breast cancer cells MCF7, and highly invasive mesenchymal breast cancer cell lines MDA-MB-436 and MDA-MB-231. Both mesenchymal cell lines with low E-cadherin expression showed very low *OVOL1* mRNA and protein expression (Supplementary Fig. 1d, e). In contrast, *OVOL1* was expressed at high levels in epithelial cell lines, in which E-cadherin was also highly expressed (Supplementary Fig. 1d, e). Notably, survival analysis revealed that breast cancer patients with higher *OVOL1* expression exhibited significantly increased relapse free survival probabilities than those with lower *OVOL1* expression (Fig. 1c)^{40, 41}. Moreover, *OVOL1* expression was lower in breast tumor samples than tumor adjacent normal-like samples (Supplementary Fig. 2a)^{39, 42}. Additional analysis suggested that *OVOL1* expression was higher in ER positive (ER+) samples than ER negative (ER-) samples with poor outcomes (Supplementary Fig. 2b)^{39, 42}. Subsequently, immunohistochemical staining results demonstrated that OVOL1 levels were reduced in invasive breast ductal carcinoma tissues compared with matched adjacent tissue samples, in 72.5 % (29 of 40) of the paired samples (Fig. 1d). Unpaired analysis also indicated that OVOL1 protein expression was decreased in carcinoma specimens (Fig. 1d). Of note, OVOL1 protein expression was substantially less in samples grouped into more advanced grades (Grade 2 and Grade 3) compared with a more benign grade (Grade 1; Fig. 1e). Taken together, our data suggest that *OVOL1* expression is inversely correlated with EMT and is a potential indicator for a favorable prognosis in breast cancer patients.

OVOL1 inhibits EMT, migration, extravasation and early phase metastatic events in breast cancer cells

Given the inverse correlation between *OVOL1* and EMP, we sought to examine the effect of OVOL1 depletion on EMT in breast cancer cells. Upon OVOL1 knockdown by two independent shRNAs in pre-malignant MCF10A-M2 breast cells, E-cadherin expression was significantly decreased, while the mesenchymal markers expression, such as Fibronectin, N-

cadherin, Vimentin and SLUG, was considerably increased (Fig. 2a and Supplementary Fig. 3a).

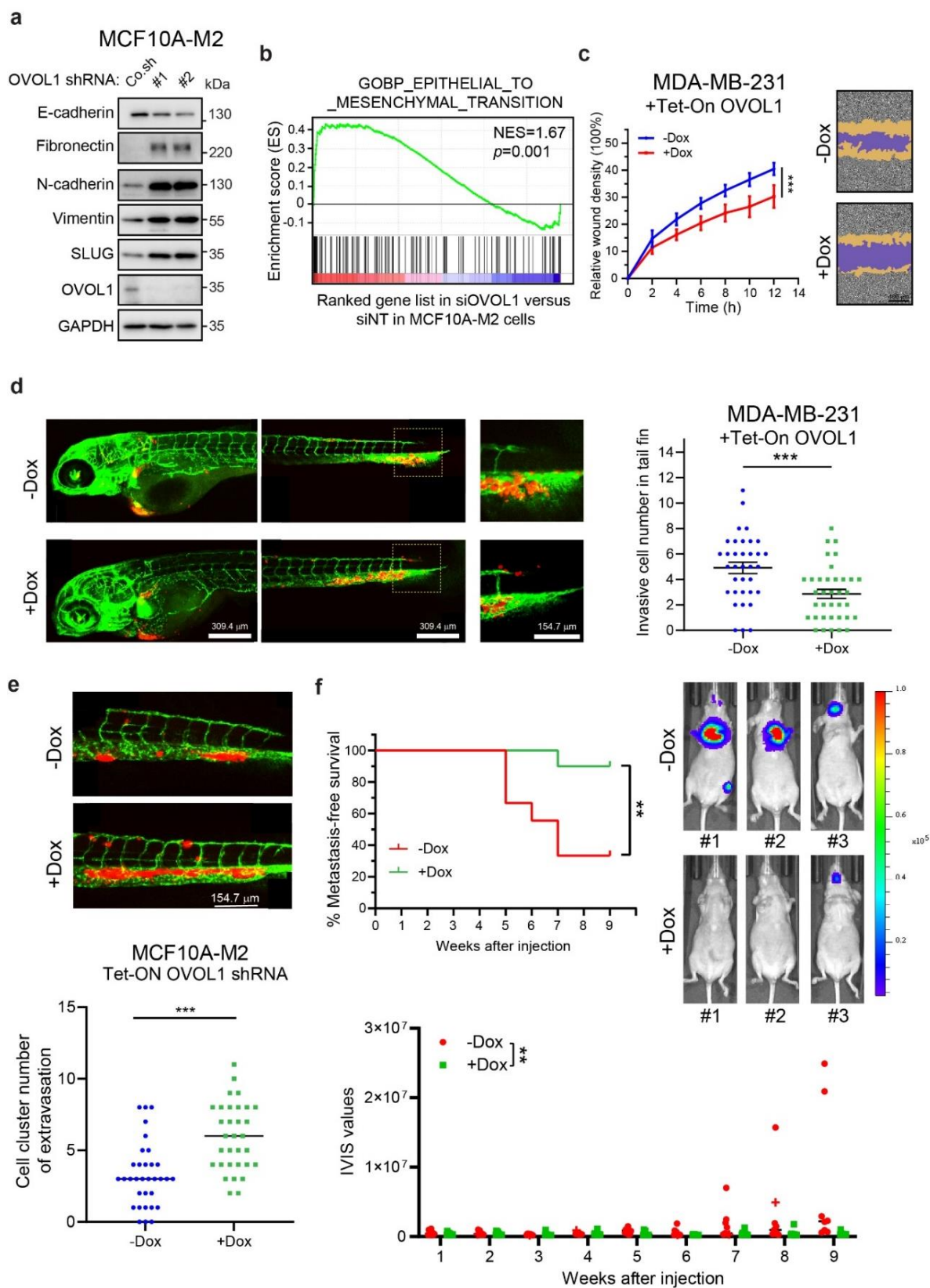


Fig. 2 OVOL1 inhibits EMT, migration, extravasation and early metastatic growth of breast cancer cells. (a) Western blotting quantification of EMT markers expression in MCF10A-M2 cells upon OVOL1 depletion. To control for equal loading

GAPDH levels were analyzed. (b) GSEA analysis of positive correlation between (manipulated) OVOL1 expression and EMT gene signature. (c) Analysis of real-time migration of MDA-MB-231 cells in the absence or presence of OVOL1 ectopic expression. Cells were either not treated or treated with Doxycycline (Dox) for 2 days prior to seeding. Relative wound density (closure) was plotted at indicated time points (left). Representative scratch wounds are shown at the end time point of the experiment (right). The regions of original scratches and the areas of migrating cells are colored in purple and yellow, respectively. Five biological replicates were included in this assay. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$. (d) *In vivo* zebrafish extravasation experiments of MDA-MB-231 cells without or with ectopic expression of OVOL1. MQ water or Dox (to enable induction of OVOL1 expression) was added to egg water from the first day post-injection. Representative images with zoom-in of the tail fin area are shown in the left panel. Analysis of the extravasated cell numbers in indicated groups is shown in the right panel. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$. (e) *In vivo* zebrafish extravasation experiments of MCF10A-M2 cells without or with the knockdown of OVOL1. MQ water or Dox (to enable induction of the shRNA targeting OVOL1) was added to the egg water from the first day post-injection. Representative images are shown in the upper panel. Analysis of the extravasated cell clusters in indicated groups is shown in the lower panel. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$. (f) Mouse xenograft cancer model of MDA-MB-231 cells without or with OVOL1 ectopic expression. MQ water or Dox (to enable induction of OVOL1 overexpression) was added to the drinking water from the first day post-injection. Metastasis-free survival is depicted in the upper left panel. Log-rank test was used for statistical analysis. ** $0.001 < p < 0.01$. Whole body bioluminescence images (BLI) at 9 weeks of mice are shown in the upper right panel. Analysis of the IVIS values in indicated groups is shown in the lower panel. One mouse in the -Dox group that was terminated at 8 weeks after injection is indicated as a cross (and not taken along in statistical analysis of BLI measurements). The results are expressed as mean \pm SD. Two-way ANOVA was used for statistical analysis. ** $0.001 < p < 0.01$.

In addition, gene set enrichment analysis (GSEA) demonstrated that OVOL1 depletion was positively correlated with the EMT gene set (Fig. 2b). Moreover, upon OVOL1 knockdown using a doxycycline (Dox)-controlled Tet-On system, epithelial HaCaT keratinocytes obtained mesenchymal-like features as analyzed based on filamentous (F)-actin staining (Supplementary Fig. 3b). The increased cell size upon OVOL1 knockdown was indeed to be expected from cells undergoing EMT. On the contrary, upon the induction of OVOL1 in MDA-MB-231 cells, in which OVOL1 expression was placed under the control of a Dox-induced Tet-On system, epithelial markers (*ECAD*, *EPCAM* and *KRT18*) expression was enhanced, while mesenchymal markers (*VIM* and *ACTA2*) expression was reduced (Supplementary Fig. 3c). Western blotting analysis also showed that OVOL1 ectopic expression resulted in an increase of E-cadherin expression and a decrease of Vimentin expression in MDA-MB-231 cells (Supplementary Fig. 3d). Wound healing and chemotaxis (transwell migration) assays revealed that OVOL1 ectopic expression led to a decrease of MDA-MB-231 cell migration (Fig. 2c and Supplementary Fig. 3e). Yet, the viability of MDA-MB-231 cells was not affected by OVOL1 overexpression (Supplementary Fig. 3f).

Next, mCherry labeled MDA-MB-231 or MCF10A-M2 cells with OVOL1 misexpression were subjected to a zebrafish embryo xenograft model⁴³. Zebrafish fed with Dox that allowed for inducing OVOL1 ectopic expression demonstrated less extravasation of MDA-MB-231 cells into the avascular tail fin area (Fig. 2d). However, MCF10A-M2 cells formed less clusters between blood vessels when zebrafish were fed with Dox to induce OVOL1 knockdown in breast cells (Fig. 2e). Next, we tested the effect of doxycycline-inducible expression of OVOL1 in MDA-MB-231 cells on their metastasis ability after intracardial injection in nude mice. The first MDA-MB-231 cell metastasis was detected much later in the mice that were fed with Dox than without Dox treatment (Fig. 2f). In addition, overexpressing OVOL1 in MDA-MB-231 cells significantly reduced the early metastatic colonization of circulating MDA-MB-231 cells (Fig. 2f, Supplementary Fig. 3g). In conclusion, our results implicate that OVOL1 mitigates the EMT, migration, extravasation and early phase metastatic events in breast cancer cells.

OVOL1 and BMP pathway form a positive feedback loop

Considering the earlier finding that OVOL1 is expressed at very low to non-detectable levels in basal breast cancer cell lines (Fig. 1b), we speculated that the promoter of *OVOL1* might be epigenetically silenced in these cells. Indeed, RT-qPCR assays showed that 5-Aza-2'-

Deoxycytidine (5-AZA), an agent for inducing DNA demethylation, was capable to upregulate *OVOL1* expression in MDA-MB-231 cells (Fig. 3a).

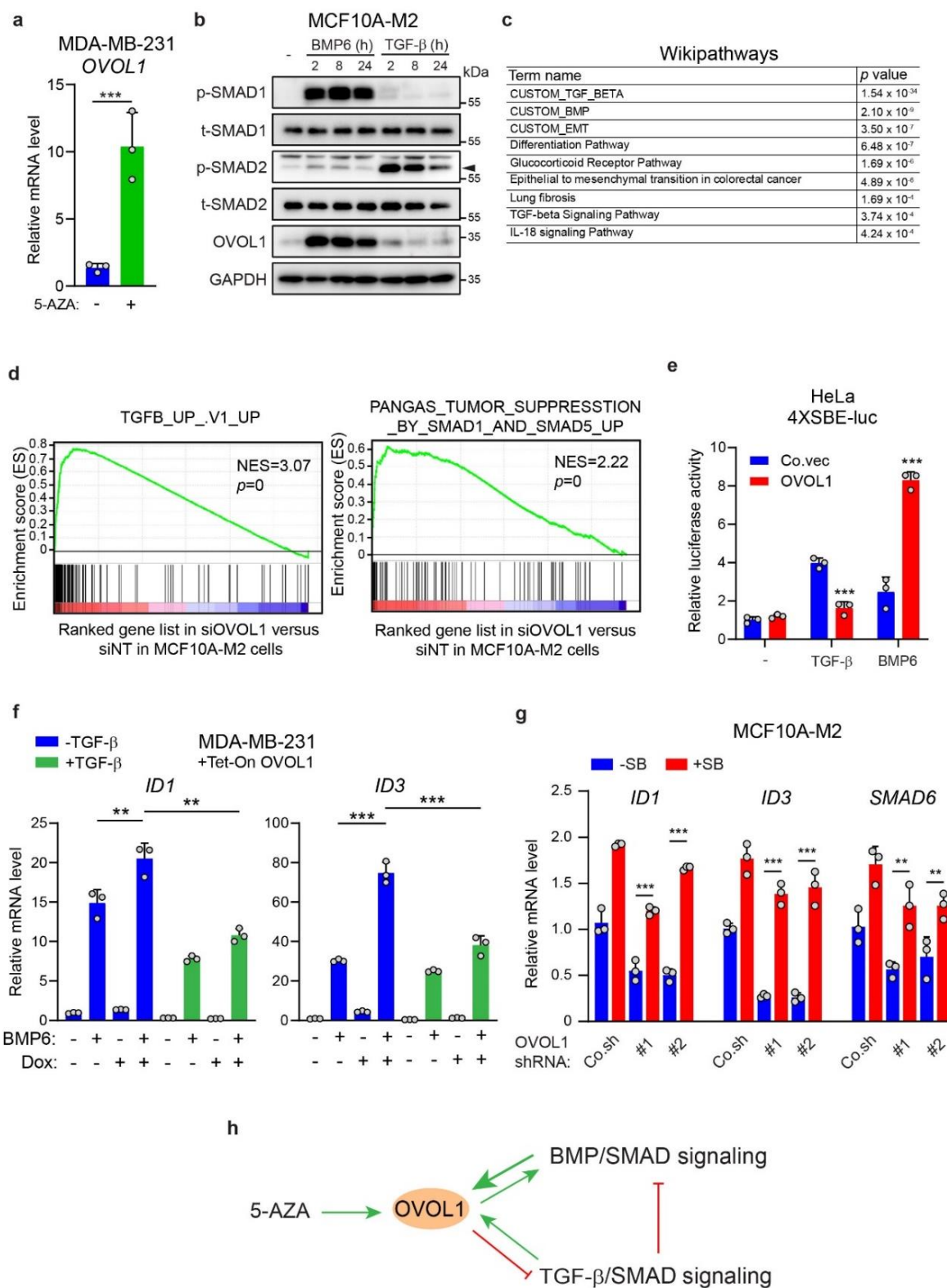


Fig. 3 *OVOL1* potentiates BMP pathway through mitigating TGF- β signaling. (a) RT-qPCR assay of *OVOL1* expression in MDA-MB-231 cells stimulated without or with 5-Aza-2'-Deoxycytidine (5-AZA) for 7 days. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$. (b) Western blotting detection of *OVOL1* expression in MCF10A-M2 cells treated with TGF- β (5 ng/ml) or BMP6 (50 ng/ml) for indicated time points. The phosphorylation of SMAD1 (p-SMAD1) or SMAD2 (p-SMAD2) was detected to confirm the activation of BMP or TGF- β pathway, respectively. To control for equal loading GAPDH

levels were analyzed. (c) The pathway enrichment results from wikipathways upon the depletion of OVOL1 in MCF10A-M2 cells. (d) GSEA analyses of the positive and inverse correlations between (manipulated) OVOL1 expression level and the TGF- β or SMAD1/5 (BMP) gene response signature, respectively. (e) Quantification of the luciferase transcriptional activity in HeLa cells transfected with TGF- β /BMP/SMAD-responsive SBE4-luc reporter and empty vector (Co.vec) or OVOL1. The results are expressed as mean \pm SD. *** 0.0001 < p < 0.001. (f) Detection of *ID1* and *ID3* expression via RT-qPCR in MDA-MB-231 cells with OVOL1 expression induced by Doxycycline (Dox). Cells were treated without or with TGF- β (5 ng/ml) for 1 day followed by Dox stimulation for 2 days. Afterwards, cells were serum starved overnight and stimulated without or with BMP6 (50 ng/ml) for 2 h. The results are expressed as mean \pm SD. ** 0.001 < p < 0.01, *** 0.0001 < p < 0.001. (g) RT-qPCR quantification of *ID1*, *ID3* and *SMAD6* levels in MDA-MB-231 cells with the inducible expression of OVOL1. Cells were either not treated or treated with SB431542 (SB) for 1 day followed by transduction of empty vector control (Co.sh) or shRNAs targeting OVOL1. The results are expressed as mean \pm SD. ** 0.001 < p < 0.01, *** 0.0001 < p < 0.001. (h) Schematic model illustrating the interplay of BMP and TGF- β pathways with OVOL1.

Given the finding that *OVOL1* is a target gene of BMP and TGF- β pathways in SMAD4-restored MDA-MB-468 cells and keratinocytes⁴⁴⁻⁴⁶, we tested if *OVOL1* can be induced by BMP and TGF- β in other breast cell lines. In MCF10A-M1 and MCF10A-M2 cells, a strong induction of *OVOL1* expression by BMP6 was observed, whereas TGF- β -mediated induction of *OVOL1* expression was weak (Fig. 3b and Supplementary Fig. 4a, b). In MDA-MB-231 cells, *OVOL1* mRNA could only be moderately induced by TGF- β and BMP6, and a kinetic response upon ligand stimulation was observed (Supplementary Fig. 4c). In addition, treating cells with selective small molecule kinase inhibitors of BMPRI (LDN193189; LDN)⁴⁷ or T β RI type I receptor (SB431542; SB)⁴⁸ efficaciously mitigated the upregulation of *OVOL1* induced by BMP6 or TGF- β (Supplementary Fig. 4d, e). To further determine whether *OVOL1* is a direct target gene of BMP/SMAD and TGF- β /SMAD pathways, we depleted SMAD4 that is essential for the transduction of both pathways^{49, 50}. BMP/TGF- β -induced *OVOL1* expression was totally blocked in the absence of SMAD4 (Supplementary Fig. 4f, g). Taken together, our data suggest that *OVOL1* is a target gene of both BMP and TGF- β signaling in breast cancer cells.

Due to the fact that a number of target genes of BMP and TGF- β , such as *SMAD6* and *SMAD7*^{51, 52}, can function as modulators to precisely control the pathway transduction, we asked whether OVOL1 regulates BMP/SMAD or TGF- β /SMAD signaling. To this end, we performed RNA-seq-based transcriptional profiling in MCF10A-M2 cells with OVOL1 depletion (Supplementary Fig. 4h). Pathway enrichment analysis revealed that TGF- β /SMAD and BMP/SMAD pathways were enriched as top pathways modulated by OVOL1 knockdown (Fig. 3c). GSEA analysis confirmed that loss of OVOL1 positively and negatively correlated with TGF- β and SMAD1/5 (BMP) gene response signatures, respectively (Fig. 3d). Yet, no TGF- β /SMAD or BMP/SMAD pathway enrichment in OVOL2 knockdown cells was revealed by GSEA analysis (Supplementary Fig. 4i, j). As indicated by a TGF- β /BMP/SMAD-driven SBE4-luc reporter⁵³, TGF- β /SMAD signaling was less active, while BMP/SMAD signaling was potentiated, in OVOL1 overexpressing HeLa cells (Fig. 3e). Nevertheless, ectopic expression of OVOL2 to a comparable level as OVOL1 did not regulate the reporter activities as potently as OVOL1 did (Supplementary Fig. 4k, l). In agreement with this result, exposing MDA-MB-231 cells to OVOL1 overexpression significantly augmented the expression of BMP target genes *ID1* and *ID3* (Supplementary Fig. 4m). As TGF- β has been reported to antagonize BMP signaling¹⁷, we presumed that the inhibitory effect of OVOL1 on TGF- β signaling might be a reason for the potentiation of BMP signaling. Consistent with this notion, the upregulation of BMP target genes upon OVOL1 ectopic expression was significantly alleviated when cells were pre-challenged with TGF- β (Fig. 3f). In contrast, blocking endogenous TGF- β signaling with a selective T β RI kinase inhibitor (SB431542; SB) rescued the reduction of BMP target genes expression imposed by the absence of OVOL1 (Fig. 3g). Collectively, these results indicate that *OVOL1* is strongly induced by BMP and mildly induced by TGF- β . In turn, OVOL1 has the capacity to augment BMP/SMAD pathway, which is

achieved (at least in part) by the inhibition, exerted by OVOL1, on TGF- β /SMAD signaling (Fig. 3h).

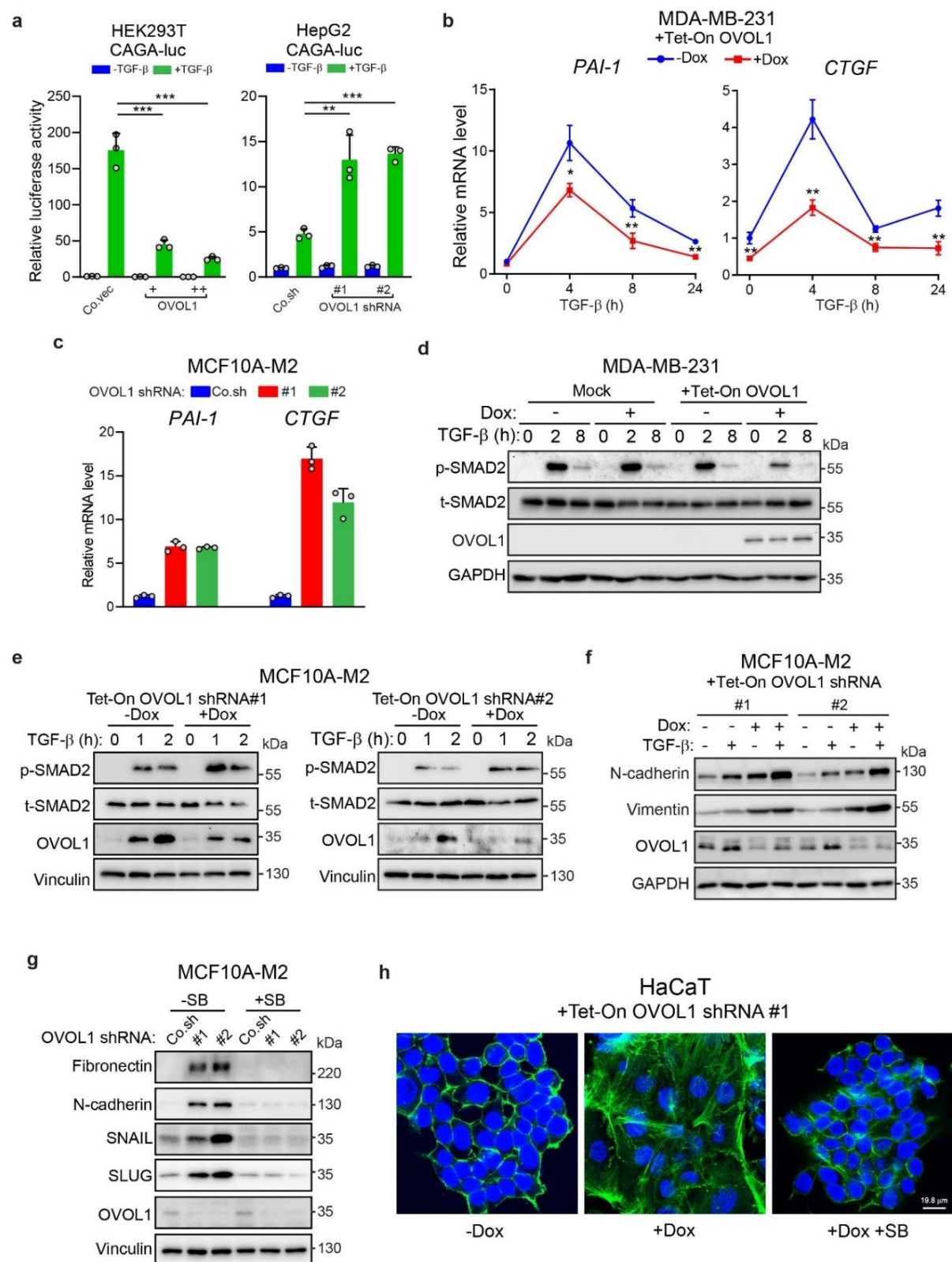


Fig. 4 OVOL1 inhibits TGF- β /SMAD signaling pathway and TGF- β -induced EMT. (a) Reporter assays for measuring the luciferase activity in HEK293T (left panel) or HepG2 cells (right panel) transfected with TGF- β -induced SMAD3/4-dependent CAGA-luc transcriptional reporter and indicated constructs for the ectopic expression (left) or depletion (right) of OVOL1. The results are expressed as mean \pm SD. ** $0.001 < p < 0.01$, *** $0.0001 < p < 0.001$. (b) *PAI-1* and *CTGF* expression as

detected by RT-qPCR in MDA-MB-231 cells with OVOL1 ectopic expression induced by Doxycycline (Dox). Cells were kept in the presence or absence of Dox for 2 days before serum starvation overnight and the treatment of TGF- β (1 ng/ml) for indicated time points. Statistical analyses were carried out at the indicated time points. The results are expressed as mean \pm SD. * 0.01 < p < 0.05, ** 0.001 < p < 0.01. (c) RT-qPCR detection of *PAI-1* and *CTGF* expression in MCF10A-M2 cells upon the knockdown of OVOL1. The results are displayed as mean \pm SD in technical triplicates. (d) Western blotting quantification of the phosphorylation of SMAD2 (p-SMAD2) and total SMAD2 (t-SMAD2) in MDA-MB-231 cells without (Mock) or with inducible OVOL1 ectopic expression (+Tet-ON OVOL1). Cells were treated without or with Doxycycline (Dox) for 2 days before serum starvation overnight and the stimulation of TGF- β (1 ng/ml) for indicated time points. To control for equal loading GAPDH levels were analyzed. (e) The phosphorylation of SMAD2 (p-SMAD2) and total SMAD2 (t-SMAD2) quantified by Western blotting in MCF10A-M2 cells with inducible OVOL1 depletion. Cells were kept in the presence or absence of Doxycycline (Dox) for 2 days before the stimulation of TGF- β (1 ng/ml) for indicated time points. (f) Western blotting analysis of changes in mesenchymal markers expression in MCF10A-M2 cells upon OVOL1 knockdown induced by Doxycycline (Dox). Cells were treated without or with Dox for 2 days before the stimulation of TGF- β (2.5 ng/ml) for another 2 days. To control for equal loading GAPDH levels were analyzed. (g) Western blotting measurement of mesenchymal markers expression in MCF10A-M2 cells with the depletion of OVOL1. Cells were either not treated or treated with SB431542 (SB) for 2 days. To control for equal loading Vinculin levels were analyzed. (h) Immunofluorescence detection of F-actin and DAPI staining of HaCaT cells upon OVOL1 knockdown induced by Doxycycline (Dox). Cells were treated without or with SB431542 (SB) for 2 h and kept in the presence or absence of Dox for 48 h.

OVOL1 suppresses TGF- β /SMAD signaling pathway and TGF- β -induced EMT

We progressed by investigating how OVOL1 mitigates TGF- β signaling transduction. The activity of TGF- β -induced SMAD3/4-driven CAGA-luc transcriptional reporter¹⁵ was inhibited in OVOL1 overexpressing HEK293T cells, and was potentiated in HepG2 and MCF7 cells with OVOL1 knockdown (Fig. 4a and Supplementary Fig. 5a). These results were further supported in MDA-MB-231 cells in which the expression of classic TGF- β target genes *PAI-1* and *CTGF* was decreased upon OVOL1 overexpression (Fig. 4b). In line with this data, depletion of OVOL1 led to a striking upregulation of TGF- β target genes (Fig. 4c and Supplementary Fig. 5b). Analyses from breast cancer patient and cell line databases also demonstrated that *OVOL1* expression was inversely correlated with the TGF- β response signature (TBRs)⁵⁴ or the levels of TGF- β /SMAD target genes (Supplementary Fig. 5c-e). Moreover, TGF- β -induced SMAD2 phosphorylation, a read-out of T β RI activity, was suppressed in MDA-MB-231 cells with OVOL1 ectopic expression (Fig. 4d). As a control, TGF- β -induced p-SMAD2 remained unaffected by Dox in parental MDA-MB-231 cells (Mock; Fig. 4d). Similar results were observed in the A549 lung adenocarcinoma cells (Supplementary Fig. 6a). On the contrary, OVOL1 depletion enhanced TGF- β -induced SMAD2 phosphorylation in MCF10A-M2 cells (Fig. 4e). Moreover, TGF- β -triggered N-cadherin and Vimentin protein expression was further potentiated in OVOL1-deficient MCF10-M2 cells, whereas TGF- β -induced changes of EMT markers expression were mitigated in A549 cells upon ectopic expression of OVOL1 (Fig. 4f and Supplementary Fig. 6b). However, pre-treating MCF10A-M2 cells with SB431542 interrupted the upregulation of mesenchymal markers expression caused by the loss of OVOL1, suggesting that the induction of EMT by OVOL1 depletion relies on the potentiation of TGF- β signaling (Fig. 4g and Supplementary Fig. 6c). Finally, immunofluorescence staining of HaCaT cells showed that F-actin rearrangement triggered by the absence OVOL1 could be rescued by SB431542 treatment (Fig. 4h). In conclusion, our results demonstrate that OVOL1 is a negative regulator of TGF- β /SMAD pathway and TGF- β -driven EMT.

OVOL1 promotes T β RI degradation by interacting with and stabilizing SMAD7

Our previous results showed that total SMAD2 levels were not changed by genetically manipulating OVOL1 expression (Fig. 4d, e). This suggested that OVOL1 may modulate TGF- β /SMAD signaling at the receptor level. To test this hypothesis, T β RI expression was evaluated upon OVOL1 misexpression. As shown in Fig. 5a, T β RI protein expression was strongly repressed in MDA-MB-231 cells overexpressing OVOL1, yet *TBRI* mRNA expression remained the same, indicating that T β RI may be post-translationally inhibited by OVOL1. To investigate T β RI protein stability, a cycloheximide (CHX)-directed protein time-course

experiment was carried out. As expected, TβRI showed a shortened half-life in MDA-MB-231 cells subjected to OVOL1 ectopic expression (Fig. 5b).

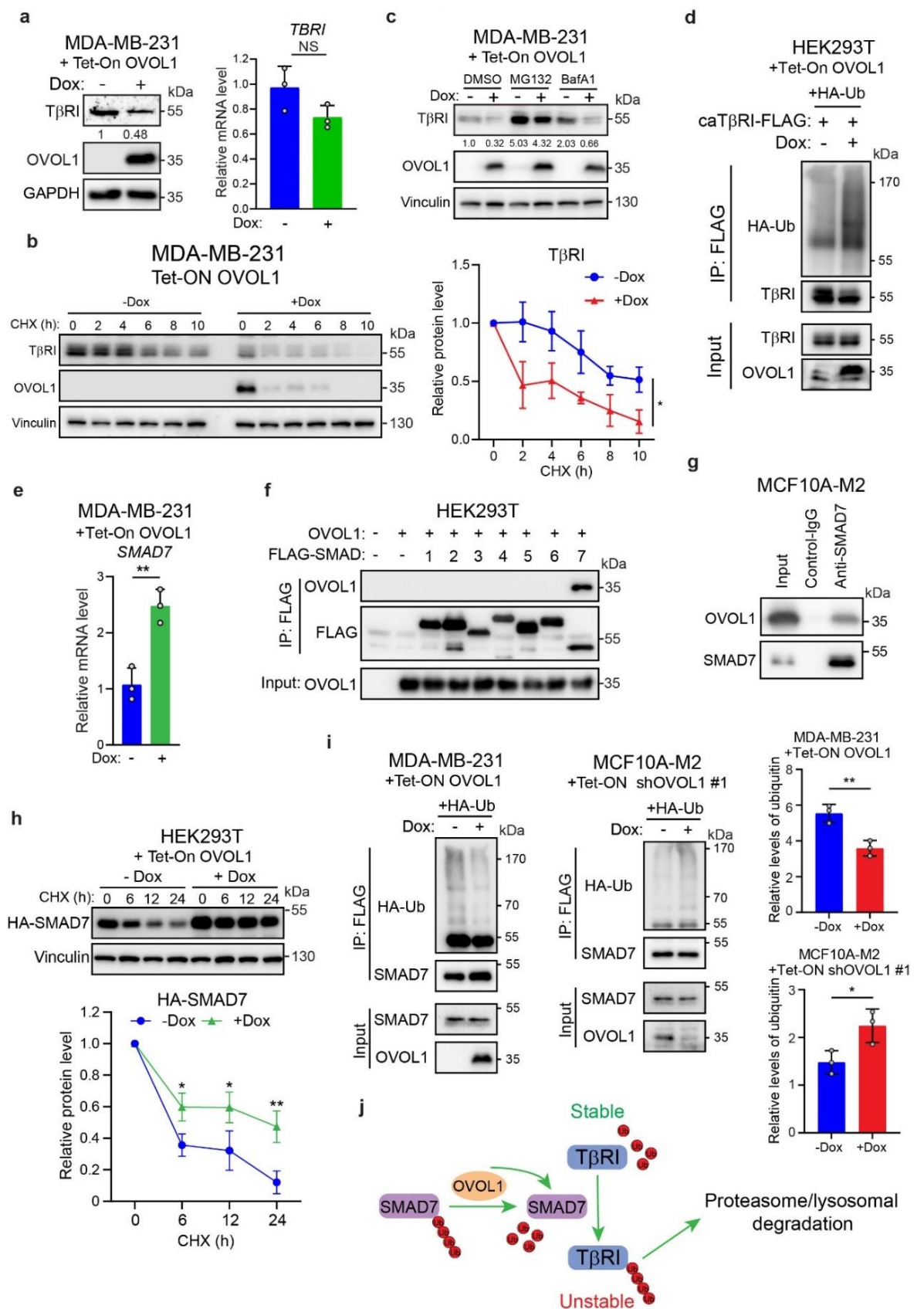


Fig. 5 OVOL1 interacts with and stabilizes SMAD7, thereby enhancing the degradation of TβRI. (a) Quantification of

TGF- β type I receptor (T β RI) protein (left) or *TBRI* mRNA expression (right) by western blotting or RT-qPCR, respectively, in MDA-MB-231 cells with inducible OVOL1 ectopic expression. Cells were kept in the presence or absence of Doxycycline (Dox) for 2 days. To control for equal loading GAPDH levels were analyzed. NS, not significant. (b) T β RI expression quantified by western blotting in MDA-MB-231 cells with OVOL1 ectopic expression induced by Doxycycline (Dox; left). Cells were cultured in the presence or absence of Dox for 2 days followed by the stimulation of cycloheximide (CHX; 50 μ g/ml) for indicated time points. Quantification of the relative protein level of T β RI is shown in the right panel. Statistical analyses were performed at the indicated time points. To control for equal loading Vinculin levels were analyzed. The results are expressed as mean \pm SD. * $0.01 < p < 0.05$. (c) Western blotting detection of T β RI expression in MDA-MB-231 cells with the inducible expression of OVOL1. Cells were kept in the presence or absence of Doxycycline (Dox) for 2 days followed by the treatment of proteasome inhibitor MG132 (5 μ M) or lysosome inhibitor BafA1 (20 nM) for 6 h. (d) Western blotting quantification of whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells with inducible OVOL1 expression. Cells were either not stimulated or stimulated with Doxycycline (Dox) for 1 day and then transfected with HA-Ub and constitutively active T β RI-FLAG (caT β RI-FLAG). (e) *SMAD7* expression measured by RT-qPCR in MDA-MB-231 cells with inducible expression of OVOL1. Cells were either not treated or treated with Doxycycline (Dox) for 2 days. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$. (f) Western blotting detection of whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells transfected with indicated FLAG-SMADs and OVOL1. (g) Western blotting quantification of 5% cell lysates (Input) and analysis of OVOL1 and SMAD7 immunoprecipitants derived from MCF10A-M2 cells. The SMAD7 antibody was added to the cell lysates to pull down SMAD7 and the IgG isotype was included as a control. (h) Western blotting measurement of the expression of HA-SMAD7 in HEK293T cells with inducible OVOL1 ectopic expression (upper). Cells were either not treated or treated with Doxycycline (Dox) for 2 days followed by the stimulation of cycloheximide (CHX; 50 μ g/ml) for indicated time points. Quantification of the relative protein level of HA-SMAD7 is shown in the lower panel. Statistical analyses were performed at the indicated time points. To control for equal loading Vinculin levels were analyzed. The results are expressed as mean \pm SD. * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$. (i) Western blotting analysis of whole cell lysates (Input) and immunoprecipitants derived from MDA-MB-231 (left) or MCF10A-M2 cells (right) stably expressing HA-Ubiquitin (HA-Ub) without or with ectopic expression of OVOL1 (left) or expression the OVOL1 targeting shRNA #1 (middle). Cells were treated without or with Doxycycline (Dox) for 2 days. Total ubiquitination of SMAD7 was probed. Quantification of the relative protein level of HA-Ubiquitin (HA-Ub) is shown in the right panel. The results are expressed as mean \pm SD. * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$. (j) Schematic working model indicating the action of OVOL1 on SMAD7 and T β RI regulation.

In contrast, depletion of OVOL1 in MCF10A-M2 cells significantly enhanced T β RI stability (Supplementary Fig. 7a). Notably, treating cells with a proteasome inhibitor MG132, but not a lysosome inhibitor BafA1, was able to bypass the OVOL1-directed destabilization of T β RI, suggesting that OVOL1 enhances T β RI degradation via the proteasomal pathway (Fig. 5c). Consistently, the polyubiquitination of T β RI was increased in OVOL1 overexpressing HEK293T cells (Fig. 5d).

Since SMAD7 plays a vital role in the proteasome-mediated ubiquitination of T β RI and OVOL1 is a well-defined transcriptional repressor³³, we considered whether *SMAD7* is transcriptionally regulated by OVOL1. Surprisingly, OVOL1 promoted the mRNA expression of *SMAD7* in MDA-MB-231 cells (Fig. 5e). As OVOL1 and SMAD7 were both localized in the nucleus (Supplementary Fig. 7b)^{24,33}, we processed by checking whether OVOL1 interacts with SMAD7 protein. Interestingly, OVOL1 was exclusively pulled down by SMAD7, indicating that SMAD7 is a potential binding partner of OVOL1 (Fig. 5f, g and Supplementary Fig. 7c). However, consistent with a previous report³⁶, OVOL2 co-precipitated with SMAD2, SMAD3 and most avidly with SMAD4, and only weakly with SMAD7 (Supplementary Fig. 7d). Next, we asked whether the protein level of SMAD7 is affected due to the interaction with OVOL1. To exclude the impact of transcription, we ectopically expressed SMAD7 in HEK293T cells using an expression plasmid in which expression was driven by a heterologous cytomegalovirus (CMV) promoter.

Western blotting analysis revealed that the SMAD7 protein expression was enhanced upon OVOL1 ectopic expression, suggesting that OVOL1 may promote SMAD7 protein stability (Supplementary Fig. 7e). This assumption was further validated by a time-course assay demonstrating that the turn-over of SMAD7 protein was significantly mitigated in HEK293T cells overexpressing OVOL1 (Fig. 5h). Furthermore, SMAD7 polyubiquitination was greatly reduced in cells subjected to OVOL1 ectopic expression (Fig. 5i and Supplementary Fig. 7f).

On the contrary, SMAD7 polyubiquitination was potentiated in the absence of OVOL1 (Fig. 5i).

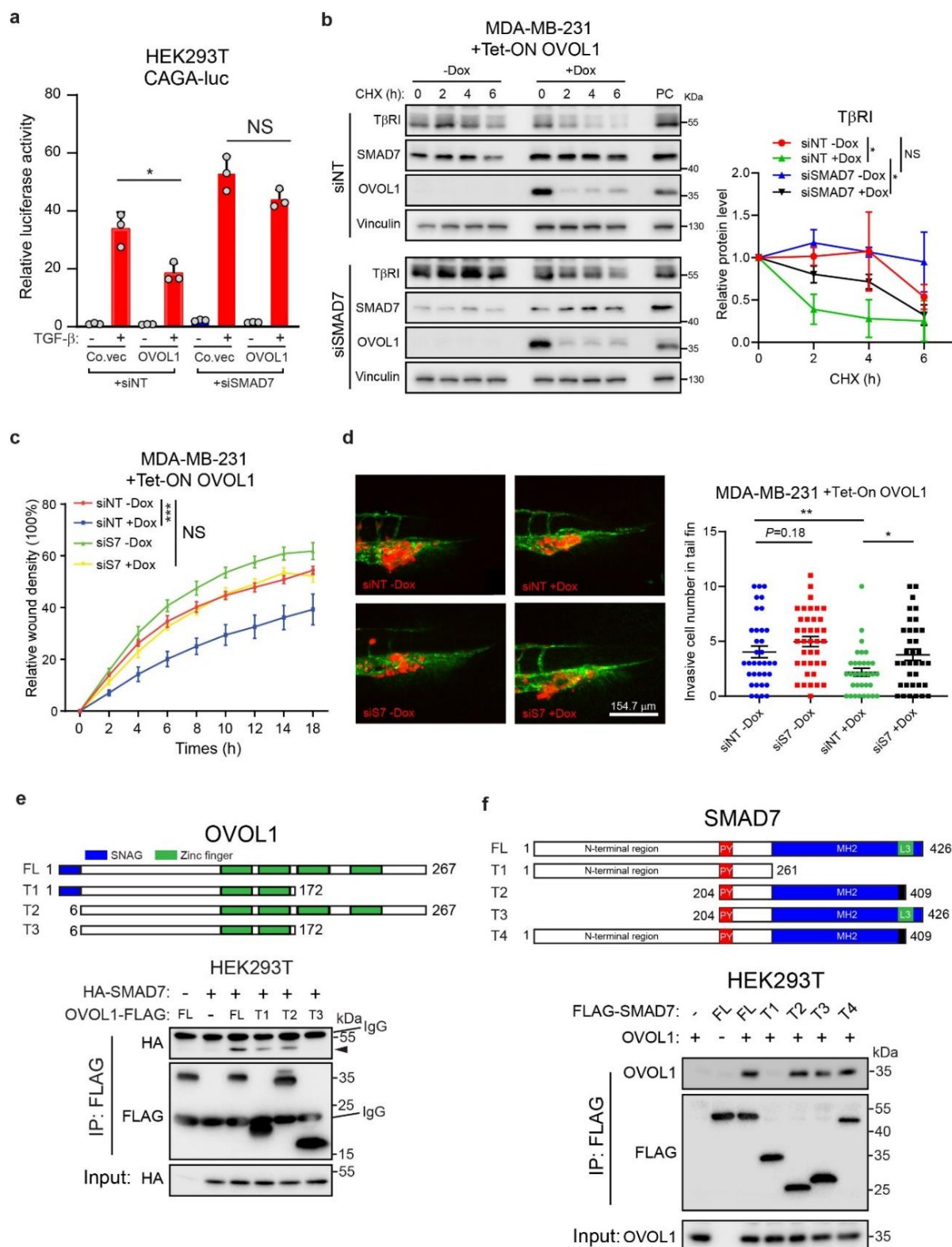


Fig. 6 The mitigation of TGF- β pathway induced by OVOL1 is SMAD7-dependent. (a) Luciferase activity in HEK293T cells transfected with TGF- β -induced SMAD3/4-dependent CAGA-luc transcriptional reporter, OVOL1 and siRNA targeting SMAD7 (siSMAD7). Non-targeting siRNA (siNT) was set as a control for siSMAD7. The results are expressed as mean \pm SD. * $0.01 < p < 0.05$. NS, not significant. (b) Western blotting detection of T β RI expression in MDA-MB-231 cells with the

inducible expression of OVOL1. Cells were transfected with siRNA targeting SMAD7 (siSMAD7) or non-targeting siRNA (siNT) for 1 day. Afterwards, cells were kept in the presence or absence of Doxycycline (Dox) for 2 days followed by the stimulation with cycloheximide (CHX; 50 μ g/ml) for indicated time points. To control for equal loading Vinculin levels were analyzed. Quantification of the relative protein levels of T β RI is shown in the right panel. The results are expressed as mean \pm SD. Statistical analysis by one-way ANOVA was carried out between indicated groups. * 0.01 < p < 0.05. NS, not significant. (c) Real-time scratch results of MDA-MB-231 cells without or with ectopic OVOL1 expression. Cells were transfected with siRNA targeting SMAD7 (siS7), followed by the treatment of Doxycycline (Dox) for 2 days prior to seeding. Non-targeting siRNA (siNT) was set as a control for siSMAD7. Relative wound density (closure) was plotted at indicated time points. Three biological replicates were included in this assay. The results are expressed as mean \pm SD. *** 0.0001 < p < 0.001. NS, not significant. (d) *In vivo* zebrafish extravasation assay of MDA-MB-231 cells without or with ectopic OVOL1 expression. Cells were transfected with non-targeting siRNA (siNT) or siRNA targeting SMAD7 (siS7), followed by the injection into zebrafish. MQ water or Doxycycline (Dox) was added to the egg water from the first day post-injection. Representative images of the tail fin area are shown in the upper panel. Analysis of the extravasated cell numbers in indicated groups is shown in the lower panel. The results are expressed as mean \pm SD. * 0.01 < p < 0.05, ** 0.001 < p < 0.01. (e) Schematic representation of the full-length OVOL1 (FL) and truncation mutants (T1-T3) tested (upper). Western blotting analysis of whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells transfected with HA-SMAD7 and FLAG tagged full-length OVOL1 or indicated OVOL1 truncation mutants is shown in the lower panel. (f) Schematic representation of full-length SMAD7 (FL) and truncation mutants (T1-T4) tested (upper panel). Western blotting analysis of whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells transfected with OVOL1 and FLAG tagged full-length SMAD7 or indicated SMAD7 truncation mutants is shown in the lower panel.

Given the fact that various ubiquitin chain topologies serve as signals to guide substrates towards different outcomes⁵⁵, we then analyzed whether OVOL1 impacts the levels of K48 or K63-incorporated ubiquitin chains on SMAD7. In particular, we identified that the K48 ubiquitin labeled SMAD7, which contributes to protein degradation, was decreased in the presence of OVOL1 overexpression (Supplementary Fig. 7g). However, the K63 ubiquitin level of SMAD7 was not affected by ectopically expressing OVOL1 (Supplementary Fig. 7g). Furthermore, we hypothesized that the interactions between SMAD7 and E3 ligases or DUBs may be altered by OVOL1. Co-immunoprecipitation experiments showed that although interactions between SMAD7 and the reported E3 ligases or DUBs for SMAD7, including ARKADIA, USP26 and OTUD1²⁷⁻²⁹, were not affected by OVOL1, the E3 ligase RNF12²⁶ and SMAD7 interaction was diminished upon the ectopic expression of OVOL1 (Supplementary Fig. 8a). As a consequence, SMAD7 polyubiquitination triggered by RNF12 was partially alleviated when OVOL1 was ectopically expressed (Supplementary Fig. 8b). These results reveal that RNF12 is involved in the decrease of SMAD7 polyubiquitination directed by OVOL1 (Supplementary Fig. 8c). Collectively, our data indicate that OVOL1 upregulates SMAD7 expression at both transcriptional and post-translational levels, which results in the polyubiquitination and degradation of T β RI (Fig. 5j).

Interaction with SMAD7 is required for the TGF- β signaling mitigation exerted by OVOL1

Next, we asked whether OVOL1 exerts its inhibitory effect on TGF- β /SMAD pathway by regulating SMAD7. Consistent with this notion, OVOL1 and SMAD7 suppressed TGF- β /SMAD-driven CAGA-luc transcriptional activity in a synergistic manner (Supplementary Fig. 9a). Conversely, the inhibition of TGF- β /SMAD downstream activity and T β RI stability by OVOL1 overexpression was alleviated upon siRNA-mediated SMAD7 knockdown (Fig. 6a, b and Supplementary Fig. 9b). In addition, OVOL1-induced attenuation of MDA-MB-231 cell migration and extravasation was rescued upon the depletion of SMAD7 (Fig. 6c, d and Supplementary Fig. 9c). Collectively, the inhibitory effects of OVOL1 on TGF- β /SMAD signaling, cell migration and extravasation are mediated (at least in part) by SMAD7. To better uncover the mechanism by which OVOL1 interacts with SMAD7, a panel of truncated mutants of OVOL1 protein were generated (Fig. 6e). Co-immunoprecipitation assays showed that the SNAG domain and last two Zinc finger domains of OVOL1 contributed to its optimum interaction with SMAD7 (Fig. 6e). To further assess the functional involvement of SMAD7 in OVOL1-mediated inhibition of TGF- β pathway, MDA-MB-231 cells were exposed to

lentivirus bearing constructs that encode full-length (FL) or truncated (T1-T3) OVOL1 proteins, respectively. Interestingly, none of the truncated OVOL1 proteins, whose interactions with SMAD7 were impaired, was capable of suppressing TGF- β -induced SMAD2 phosphorylation, SMAD3/4-driven CAGA-luc transcriptional activity and the induction of TGF- β target genes, as potent as the full-length OVOL1 did (Supplementary Fig. 9d-f). Moreover, the binding regions of OVOL1 on SMAD7 were mapped to further decipher their interaction. We observed that the Mad Homology-2 (MH2) domain-deficient SMAD7 failed to be co-immunoprecipitated together with OVOL1, implying that the MH2 domain of SMAD7 is enrolled in its interaction with OVOL1 (Fig. 6f).

FICZ activates OVOL1 and attenuates TGF- β pathway, EMT, cell migration and extravasation

Since OVOL1 was identified as a suppressor of TGF- β /SMAD signaling and breast cancer progression, we focused on restoring OVOL1 expression by candidate small molecule compounds. FICZ was reported to promote OVOL1 expression in keratinocytes⁵⁶, which prompts us to investigate if it is eligible to generate the similar effect in pre-malignant breast cells. It is noteworthy that FICZ substantially upregulated *OVOL1* mRNA and protein levels in MCF10A-M2 cells (Fig. 7a, b). These results suggested that this drug may be employed to reactivate OVOL1 expression and thereby repressing TGF- β -induced signaling in pre-malignant cancer cells. Indeed, we observed that TGF- β -induced p-SMAD2, T β RI stability and SMAD7 polyubiquitination were impinged when MCF10A-M2 cells were stimulated with FICZ (Fig. 7c-e and Supplementary Fig. 10a). However, OVOL1 depletion was able to partially compensate for these inhibitory effects exerted by FICZ, suggesting that OVOL1 may contribute to the suppressive role of FICZ in TGF- β signaling (Fig. 7c-e and Supplementary Fig. 10a). Follow-up experiments showed that FICZ-induced downregulation of mesenchymal markers and inhibition of cell migration and extravasation were mitigated, to some extent, by OVOL1 knockdown (Fig. 7f-h and Supplementary Fig. 10b-d). All these data suggest that restoration of OVOL1 by FICZ offers the prospect for therapeutic gain to mitigate overactive pro-oncogenic TGF- β signaling and breast cancer progression.

Discussion

Our study showed the inhibitory effects of OVOL1 on EMT, cell migration, extravasation and the early metastatic colonization of breast cells in a mouse cancer xenograft model. We unraveled a novel mechanism by which OVOL1 attenuates TGF- β /SMAD signaling and maintains the epithelial identity of breast cancer cells (Fig. 3h). OVOL1 does so by interacting with and preventing the ubiquitination and degradation of inhibitory SMAD7 (Fig. 5h and Supplementary Fig. 8c). The increased T β RI levels mediate enhanced TGF- β /SMAD signaling and drive breast epithelial cells to transit into cells with more mesenchymal characteristics. In contrast, BMP/SMAD pathway and a small molecule compound FICZ balance the cell plasticity towards a more epithelial-like status, by (at least in part) the induction of *OVOL1* expression (Fig. 7i).

Despite a variety of evidence from experimental and mathematical modeling analyses characterize OVOL1 as an MET inducer in multiple cancers, the association between OVOL1 and EMP in breast cancer patient remains ill-defined^{35, 57-60}. Through analyzing transcriptome datasets, a negative correlation between *OVOL1* and EMT signature was uncovered in various patient cohorts (Fig. 1a). Significantly, we are the first to report that OVOL1 protein is lower expressed in breast invasive carcinoma and is inversely correlated with the progression of breast cancer towards aggressive grades, which demonstrates that the absence of OVOL1 expression may be expected to aid detecting aggressive breast cancers (Fig. 1d, e).

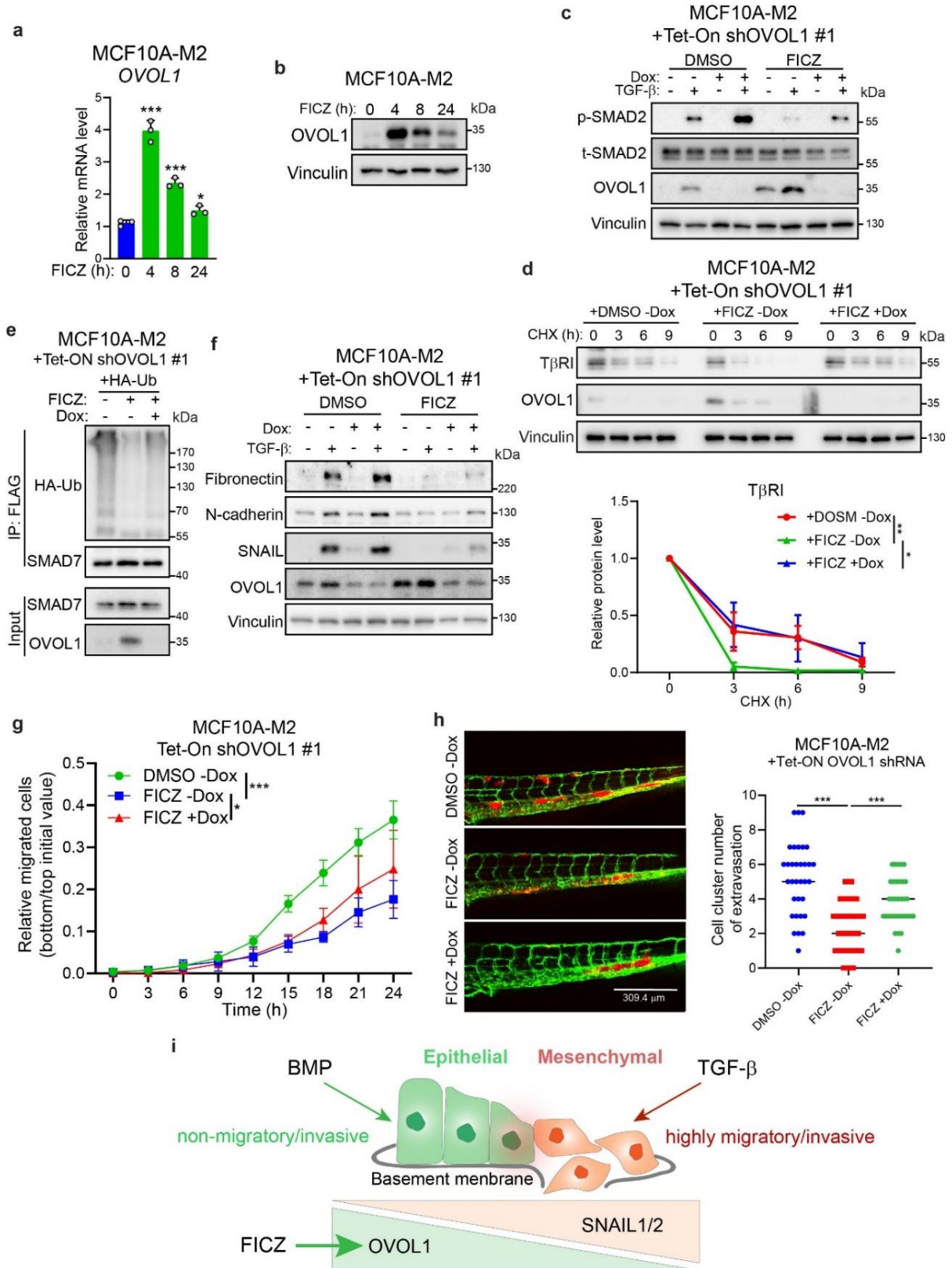


Fig. 7 FICZ upregulates OVOL1 expression and mitigates TGF- β pathway, EMT, cell migration and extravasation. (a) RT-qPCR detection of *OVOL1* expression in MCF10A-M2 cells stimulated with 6-Formylindolo(3,2-b)carbazole (FICZ; 5 μ M) for indicated time points. Statistical analyses were carried out between 0 h group and groups at indicated time points. The results are expressed as mean \pm SD. ** 0.001 < p < 0.01, *** 0.0001 < p < 0.001. (b) Western blotting quantification of OVOL1 expression in MCF10A-M2 cells stimulated with FICZ (5 μ M) for indicated time points. Vinculin levels were analyzed to control for equal loading. (c) Western blotting measurement of the phosphorylation of SMAD2 (p-SMAD2) and OVOL1 expression in MCF10A-M2 cells with inducible OVOL1 knockdown by shRNA #1. Cells were stimulated without or with

Doxycycline (Dox) for 2 days, followed by FICZ (5 μ M) treatment in serum starvation overnight before adding TGF- β (1 ng/ml) for another 2 h. The vehicle control DMSO was included for FICZ. Vinculin levels were analyzed to control for equal loading. (d) Western blotting measurement of the expression of T β RI in MCF10A-M2 cells upon OVOL1 knockdown induced by Doxycycline (Dox). Cells were either not treated or treated with Doxycycline (Dox) for 2 days followed by the stimulation of FICZ (5 μ M) overnight. Cycloheximide (CHX; 50 μ g/ml) was then added to the medium for indicated time points. Quantification of the relative protein level of T β RI is shown in the lower panel. To control for equal loading Vinculin levels were analyzed. The results are expressed as mean \pm SD. Statistical analyses were performed using one-way ANOVA. * 0.01 < p < 0.05, ** 0.001 < p < 0.01. (e) Western blotting analysis of whole cell lysates (Input) and immunoprecipitants derived from MCF10A-M2 cells stably expressing HA-Ubiquitin (HA-Ub) without or with expressing the OVOL1 targeting shRNA #1. Cells were treated without or with Doxycycline (Dox) for 2 days and FICZ (5 μ M) overnight. Total ubiquitination of SMAD7 was probed. (f) Western blotting analysis of mesenchymal markers expression in MCF10A-M2 cells with inducible OVOL1 knockdown by shRNA #1. Cells were stimulated without or with Doxycycline (Dox) for 2 days, followed by FICZ (5 μ M) treatment for 8 h before adding TGF- β (5 ng/ml) for 2 days. The vehicle control DMSO was included for FICZ. (g) IncuCyte real-time chemotaxis assay for evaluating the migration of MCF10A-M2 cells with inducible OVOL1 knockdown by shRNA #1. Cells were pre-treated without or with Doxycycline (Dox) for 2 days before being seeded into the inserts, followed by FICZ (5 μ M) treatment. The vehicle control DMSO was included for FICZ. The results are expressed as mean \pm SD. * 0.01 < p < 0.05, *** 0.0001 < p < 0.001. (h) *In vivo* zebrafish embryo xenograft extravasation experiments of MCF10A-M2 cells without or with the knockdown of OVOL1. MQ water or Dox (to enable induction of the shRNA targeting OVOL1) and DMSO or FICZ (1 μ M) was added to the egg water from the first day post-injection. Representative images are shown in the left panel. Analysis of the extravasated cell clusters in indicated groups is shown in the right panel. The results are expressed as mean \pm SD. *** 0.0001 < p < 0.001. (i) Schematic model of balancing EMP of breast cells by TGF- β , BMP and downstream transcription factors OVOL1 as well as SNAIL1/2.

The initiation of cancer metastasis involves in the augmented migratory and invasive capacities that are conferred by EMT on primary tumor cells³. Upon extravasating into the parenchyma of distant organs, it seems that cancer cells undergo MET to support the outgrowth of micrometastases^{4,61}. We have clearly shown that OVOL1 suppresses EMT, cell migration, *in vivo* extravasation using a zebrafish embryo xenograft model, and early stage metastatic colonization using a mouse xenograft model (Fig. 2a-f). Of note, we found that ectopic expression of OVOL1 does not affect MDA-MB-231 cell proliferation/viability (Supplementary Fig. 3f). Yet, we cannot exclude the possibility that OVOL1-mediated MET may potentiate the outgrowth of micrometastases at the last step of cancer metastasis *in vivo*. Therefore, further investigation is required to comprehensively examine the role of OVOL1 in the invasion-metastasis cascade⁴.

One of the key molecular hallmarks of EMT is the absence of E-cadherin expression. In highly invasive breast cancer cells including MDA-MB-231, *ECAD* promoter is hypermethylated, resulting in the loss of E-cadherin expression⁶². *ECAD* expression was shown to be dramatically enhanced in MDA-MB-231 cells ectopically expressing OVOL1, which is consistent with the data shown by Roca and colleagues³⁵, indicating that OVOL1 may be engaged in regulating the methylation of *ECAD* promoter (Supplementary Fig. 3c, d). Thus it is worth to further investigate whether OVOL1 can work together with some demethylating enzymes to alleviate the methylation status of *ECAD* promoter. Moreover, the possibility that OVOL1 may antagonize the effects of particular methylating enzymes on *ECAD* promoter cannot be excluded.

OVOL2 promoter is hypermethylated in late stage colorectal cancer patients⁶³. We observed that 5-AZA dramatically upregulated *OVOLI* mRNA expression in MDA-MB-231 cells (Fig. 3a). To determine whether this 5-AZA effect is direct or indirect, further evidence on the methylation status within the *OVOLI* promoter is needed. It has been reported that OVOL1 functions downstream of WNT signaling in differentiating epidermal cells and hair follicles⁶⁴. Here, we have identified *OVOLI* as a target gene of BMP/SMAD and TGF- β /SMAD pathways (Fig. 3b). This result is in agreement with the previous report that *OVOLI* expression can be induced by TGF- β in keratinocytes^{45, 46} or by TGF- β and BMP in MDA-MB-468 cells with SMAD4 restoration⁴⁴, although no biological relevance was investigated in these studies. In the present study, we found OVOL1 as a potentiator of BMP signaling, which is, at least

partially, achieved by the attenuation of TGF- β signaling (Fig. 3f-h). Similarly, OVOL1 was reported to enhance osteoblast differentiation by activating *BMP2* transcription⁶⁵. In contrast, the transcription of a typical BMP target gene *ID1* can be directly suppressed by OVOL1 during trophoblasts development⁶⁶. These findings indicate that the mechanism by which OVOL1 modulates BMP signaling may be context-dependent.

Given the pivotal role in mediating the cross talk between TGF- β pathway and many other signaling pathways, SMAD7 itself is under strict control by multiple mechanisms and layers of regulation²⁴. As a target gene of TGF- β /SMAD pathway and thereby mediating important negative feedback control, transcription of *SMAD7* is activated by the SMAD complex together with co-regulators²⁴. Our results indicate that OVOL1 is capable to promote SMAD7 expression at both transcriptional and post-translational levels (Fig. 5). Interestingly, although ectopic expression of OVOL1 in MDA-MB-231 cells significantly decreases the TGF- β downstream reporter activity and the expression of TGF- β target genes, i.e. *PAI-1* and *CTGF*, the *SMAD7* mRNA is strikingly increased (Fig. 5e). OVOL1 was initially identified as, and generally functions as, a transcriptional repressor by binding to the CCGTTA element within the promoter of target genes³³. The transcriptional repressor SNAIL, which also contains the N-terminal SNAG domain as OVOL proteins do, was shown to function as a transcriptional activator by binding to specific enhancer regions⁶⁷. Although the canonical binding element of OVOL1 cannot be found in the promoter region of *SMAD7* (data not shown), we cannot exclude the possibility that, upon forming a complex with other protein partners such as transcriptional activators, OVOL1 may bind to and potentiate the activity of some enhancer regions, thereby enhancing the transcription of a particular panel of genes such as *SMAD7*. Moreover, it is also likely that OVOL1 regulates the transcription of *SMAD7* in an indirect manner.

At the post-translational level, SMAD7 is targeted for polyubiquitination and degradation by E3 ligases such as ARKADIA and RNF12^{25, 26}, whose effects can be counteracted by DUBs like USP26 and OTUD1^{27, 28}. OVOL1 was observed to decrease K48-ubiquitin of SMAD7 and thereby enhancing SMAD7 protein stability (Fig. 5h-i and Supplementary Fig. 7f, g). We hypothesized that OVOL1 may act as a scaffold to bring DUBs to SMAD7 or impair interactions between E3 ligases and SMAD7. Mechanistically, we showed that RNF12 and SMAD7 interact, and that the resulting polyubiquitination of SMAD7, were attenuated by OVOL1 ectopic expression (Supplementary Fig. 8a-c). However, we cannot rule out the possibility that other proteins (DUBs or E3 ligases) may also participate in OVOL1-mediated deubiquitylation of SMAD7. To further investigate (other potential) underlying mechanisms, proteomic interactome analysis for OVOL1 can be performed to identify novel and relevant interacting E3 ligases or DUBs.

In a previous study, ectopic expression of OVOL2 was observed to inhibit TGF- β /SMAD signaling in NMuMG cells through directly interacting with SMAD4 and interfering in the complex formation between SMAD2/3 and SMAD4³⁶. Results from our SMAD-driven transcriptional (SBE) reporter assays indicated that the induction of TGF- β /SMAD signaling and BMP/SMAD signaling are more potently affected upon the ectopic expression of OVOL1 than OVOL2 (Fig. 3e and Supplementary Fig. 4k, l). Moreover, RNA-seq-based profiling of transcriptional changes and GSEA upon OVOL2 knockdown demonstrated that TGF- β /SMAD and BMP/SMAD signaling pathways are not affected (Supplementary Fig. 4i, j). OVOL1 strongly interacts with SMAD7 but not with the other SMAD proteins. OVOL2, however, interacts with SMAD2, SMAD3 and most avidly with SMAD4, but only weakly interacts with SMAD7 (Supplementary Fig. 7d). These latter results indicate that despite the structural

similarity, and in some cases functional redundancy³², these two OVOL proteins contribute to the regulation of TGF- β /SMAD (and BMP/SMAD) signaling via diverse mechanisms.

TGF- β pathway is hyperactivated during breast cancer progression, making it an attractive therapeutic target⁶⁸. We have validated the inhibitory role of OVOL1 on TGF- β pathway transduction and TGF- β -induced EMT (Fig. 4), suggesting that restoration of OVOL1 expression may be an option for targeting the pro-oncogenic TGF- β signaling in breast cancer. Importantly, we have identified a small molecule compound FICZ as an activator to trigger OVOL1 expression and thereby suppressing TGF- β /SMAD signaling, EMT, cell migration and extravasation, which may provide opportunities to mitigate breast cancer progression (Fig. 7). Although depletion of OVOL1 rescues, to some extent, the inhibitory effects of FICZ, other targets of FICZ may also contribute to its suppressive role on breast cancer progression. FICZ is an UV-derived tryptophan photoproduct whose inhibitory effects have been shown in inflammatory diseases such as chronic mite-induced dermatitis⁶⁹. It will be interesting to explore the potential of FICZ in treatment of breast cancer in (pre)clinical models. Although adult mice with FICZ treatment in relatively high concentrations do not have pathological signs^{70, 71}, systematic toxicity studies in animal models are required to evaluate the clinical potential of FICZ. Moreover, OVOL1 agonists with higher specificity and *in vivo* safety can be explored to enable the therapeutic gain for breast cancer patients.

Materials and Methods

Data analyses of gene expression in clinical patient samples and breast cancer cell lines

Four publicly available gene expression datasets (GSE102484⁷², GSE36771⁷³, GSE12276⁷⁴ and GSE3494⁷⁵) derived from the R2 Genomics Analysis and Visualization Platform (R2 platform; <http://r2.amc.nl>) were applied for the analyses of the correlations between *OVOL1* mRNA levels and the EMT or TGF- β response gene signature. Using the GSE12777 dataset⁷⁶, *OVOL1*, *OVOL2* or *OVOL3* expression in 51 breast cell lines was compared. In the same dataset, correlations between *OVOL1* and EMT markers or TGF- β target genes expression were carried out. Differential expression data of *OVOL1* mRNA level in TCGA patients were derived from the MOBCdb database⁴². Correlations between *OVOL1* and the TGF- β target genes were evaluated in 1097 TCGA patients from R2 platform³⁹. Pearson's correlation coefficient tests were performed to assess the statistical significance.

Cell culture and reagents

HEK293T, HeLa, HepG2, A549, MDA-MB-231, MCF7 were purchased from American Type Culture Collection (ATCC) and SUM149PT cells were obtained from Dr. Sylvia Le Dévédec (Leiden Academic Center for Drug Research, Leiden, the Netherlands). All the cell lines were cultured in Dulbecco's modified Eagle medium (DMEM; Thermo Fisher Scientific; Cat. Nr.: 41965062) supplemented with 10% fetal bovine serum (FBS; Thermo Fisher Scientific; Cat. Nr.: 16000044) and 100 U/ml penicillin/streptomycin (Thermo Fisher Scientific; Cat. Nr.: 15140163). MCF10A-Ras (MCF10A-M2) cells were derived from MCF10A cells transformed with Ha-Ras and were kindly provided by Dr. Fred Miller (Barbara Ann Karmanos Cancer Institute, Detroit, USA). MCF10A and MCF10A-M2 cells were maintained in DMEM/F12 (GlutaMAX™ Supplement; Thermo Fisher Scientific; Cat. Nr.: 31331028) containing 5% horse serum (Thermo Fisher Scientific; Cat. Nr.: 26050088), 0.1 μ g/ml Cholera toxin (Sigma-Aldrich; Cat. Nr.: C8052), 0.02 μ g/ml Epidermal Growth Factor (EGF; Sigma-Aldrich; Cat. Nr.: 01-107), 0.5 μ g/ml Hydrocortisone (Sigma-Aldrich; Cat. Nr.: H0135), 10 μ g/ml Insulin (Sigma-Aldrich; Cat. Nr.: I6634) and 100 U/ml penicillin/streptomycin. All the cell lines mentioned were maintained in a 5% CO₂, 37 °C, humidified incubator and tested negative for mycoplasma routinely. Human cell lines were checked for authenticity by short tandem repeats

(STR) profiling. The 20 human breast cancer cell lines used for detecting OVOL1 (and epithelial and mesenchymal markers) expression were described previously⁷⁷. Doxycycline (Dox; Sigma-Aldrich; Cat. Nr.: D9891) or 5-Aza-2'-Deoxycytidine (5-AZA; Sigma-Aldrich; Cat. Nr.: 189826) was added to the culture medium at a final concentration of 100 ng/ml and 5 μ M, respectively. 6-Formylindolo(3,2-b)carbazole (FICZ; Sigma-Aldrich; Cat. Nr.: SML1489) dissolved in DMSO were used at 5 μ M. Lysosome inhibitor BafA1 (Sigma-Aldrich; Cat. Nr.: B1793) was used to treat cells at 20 nM final concentration. Selective small molecule kinase inhibitors of BMPR1 (LDN193189; LDN)⁴⁷ and T β R1 type I receptor (SB431542; SB)⁴⁸ were used at a concentration of 5 μ M and 120 nM, respectively. Recombinant BMP6 and TGF- β 3 were a kind gift from Slobodan Vukicevic (University of Zagreb) and Andrew Hinck (University of Pittsburg), respectively.

Generation of constructs

Human *OVOL1* cDNA was amplified by PCR from MCF10A-M2 cells and subcloned into the lentiviral vector pLV-bc-CMV-puro. The inducible vector for OVOL1 ectopic expression was generated using Gateway cloning into the pLIX-403 vector (Addgene; Cat. Nr.: 41395). We used two shRNAs TRCN0000015679 (#1) and TRCN0000015681 (#2) from Sigma MISSION® shRNA library for OVOL1 knockdown. The sequences targeting OVOL1 AGTGTCACAACGACGTCAAGA (#1) and AGGATTTGATGGCTACCAAAT (#2) were cloned into the lentiviral FH1tUTG vector to generate lentiviral constructs for inducible OVOL1 knockdown.

Lentiviral transduction and transfections

Third-generation lentiviral packaging vectors (VSV, gag and Rev) and cDNA or shRNA expressing constructs were transfected into HEK293T cells. Cell supernatants were collected at 48 h post-transfection. To generate stable cell lines, cells were plated at 10% confluence and infected by lentiviral supernatants supplemented with the same volume of fresh medium and 8 ng/ml Polybrene (Sigma-Aldrich; Cat. Nr.: 107689) for 24 h. After 48 h infection, cells were selected with Puromycine (1 μ g/ml; Sigma-Aldrich; Cat. Nr.: P9620) for 3 days. For transfection of non-targeting siRNA (Dharmacon), siRNA SMARTpool targeting *SMAD7* (Dharmacon; Cat. Nr.: L-020068-00-0005), *OVOL1* (Dharmacon; Cat. Nr.: L-006543-01-0005) or *OVOL2* (Dharmacon; Cat. Nr.: L-013793-02-0005), cells were seeded at 80% confluence and incubated with complex formed by DharmaFECT transfection reagents and siRNA (10 nM at final concentration). Medium was changed at 24 h post-transfection. RNA samples were collected at 2 days after transfection.

Real-time quantitative PCR (RT-qPCR)

RNA was extracted by a NucleoSpin RNA kit (Macherey Nagel; Cat. Nr.: 740955) according to the manufacturer's instructions. Subsequently, reverse transcription was performed using a RevertAid RT Reverse Transcription Kit (Thermo Fisher Scientific; Cat. Nr.: K1691). Indicated genes were detected by specific primer pairs on the generated cDNA using the CFX Connect Real-Time PCR Detection System (Bio-Rad). *GAPDH* was used as a reference transcript. The results are expressed as mean \pm SD, n = 3. $2^{-\Delta\Delta C_t}$ method was applied to analyze the relative expression. Primer sequences used in this study are listed in Supplementary Table 1.

RNA-seq-based transcriptional profiling, pathway enrichment and GSEA analysis

Two days post-transfection, cells were collected for RNA samples preparation by the NucleoSpin RNA kit. After mRNA enrichment by Oliogo dT selection and the following library preparation, RNA-seq was performed in the DNBseq platform (Beijing Genomics Institute, BGI, Hongkong). Afterwards, RNA-Seq files were processed using the opensource BIODDL

RNAseq pipeline v5.0.0 (<https://zenodo.org/record/5109461#.Ya2yLFPMJhE>) developed at the LUMC. This pipeline performs FASTQ preprocessing (including quality control, quality trimming, and adapter clipping), RNA-Seq alignment, read quantification, and optionally transcript assembly. FastQC was used for checking raw read quality. Adapter clipping was performed using Cutadapt (v2.10) with default settings. The alignment of RNA-Seq reads was carried out using STAR (v2.7.5a) on GRCh38 human reference genome. The gene read quantification was conducted using HTSeq-count (v0.12.4) with setting “–stranded=no”. The gene annotation used for quantification was Ensembl version 104. Using the gene read count matrix, counts per million mapped reads (CPM) was calculated per sample on all annotated genes. Genes with a higher CPM than 1 in at least 25% of all samples are kept for downstream analysis. This provided us with 11574 genes for the analysis between siOVOL1 and siNT group, and 11565 genes for the analysis between siOVOL2 and siNT group. For the differential gene expression analysis, dgeAnalysis R-shiny application (<https://github.com/LUMC/dgeAnalysis/tree/v1.4.4>) was used. EdgeR (v3.34.1) with trimmed mean of m-values (TMM) normalization was used to perform differential gene expression analysis. Benjamini and Hochberg false discovery rate (FDR) was computed to adjust p-values obtained for each differentially expressed gene. Using a cutoff of 0.05 at the adjusted p-values, we identified all up and down regulated genes. Using the differentially expressed genes as inputs, the pathway enrichment analysis was then performed with the aid of wikipathways symbol in gProfiler R package. Gene set enrichment analysis (GSEA) was carried out using the GSEA software^{78,79}. TGF- β (TGFB_UP.V1_UP) gene response signature⁸⁰, SMAD1/5 (BMP) (PANGAS_TUMOR_SUPPRESSION_BY_SMAD1_AND_SMAD5_UP) gene response signature⁸¹ and EMT (GOBP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION; GO: 0001837) gene signature were set as references to determine the correlations between (manipulated) OVOL1 or OVOL2 expression and TGF- β /SMAD signaling, BMP/SMAD signaling or EMT, respectively.

Western blotting

Cells were lysed with RIPA buffer (150 mM Sodium Chloride, 1.0% Triton X-100, 0.5% Sodium Deoxycholate, 0.1% SDS and 50 mM Tris-HCl, pH 8.0) containing freshly added complete protease inhibitor cocktail (Roche; Cat. Nr.: 11836153001). Protein concentrations were measured by a DCTM protein assay kit (Bio-Rad; Cat. Nr.: 5000111) according to the manufacturer's instructions. Equal amounts of proteins were loaded and separated by Sodium Dodecyl Sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Afterwards, proteins were transferred onto a 45- μ m polyvinylidene difluoride (PVDF) membrane (Merck Millipore; Cat. Nr.: IPVH00010). 5% non-fat dry milk dissolved in Tris-buffered saline (TBS) with 0.1% Tween 20 (TBST) was used to block the membrane for 1 h at room temperature (RT). Membranes were probed with the respective primary and secondary antibodies. ClarityTM Western ECL Substrate (Bio-Rad; Cat. Nr.: 1705060) and ChemiDoc Imaging System (Bio-Rad; Cat. Nr.: 17001402) were used to detect the signal. Primary antibodies used in this study are listed in Supplementary Table 2. Secondary antibodies used in this study are anti-IgG (Sigma-Aldrich; Cat. Nr.: NA931V) and anti-rabbit (Cell Signaling; Cat. Nr.: 7074S). All results were derived from at least three independent biological replicates, and representative results are shown. Protein levels were quantified by densitometry using ImageJ (National Institutes of Health, United States).

Co-immunoprecipitation assays

HEK293T or MCF10A-M2 cells were lysed with TNE lysis buffer (50 mM Tris-HCl, pH 7.4, 1 mM EDTA, 150 mM NaCl and 1% NP40) containing freshly added complete protease inhibitor cocktail and kept on ice for 15 min. The lysates were centrifuged at 1.4×10^4 g for 10

min at 4°C. Equal amounts of protein were incubated with anti-FLAG agarose beads (Sigma-Aldrich; Cat. Nr.: A2220) for 30 min at 4°C with rotation. For checking the interaction between endogenous SMAD7 and OVOL1, 1 μ L antibodies against SMAD7 (R&D; Cat. Nr.: MAB2029) were added to the cell lysates and the mixtures were kept overnight at 4°C with rotation. The next day, 20 μ L Pierce™ Protein G Agarose (Thermo Fisher; Cat. Nr.: 20397) were added and kept for 2 h at 4°C with rotation. Beads were washed for five times with the TNE buffer for 5 min at 4°C with rotation. Afterwards, samples were boiled with 2 \times sample buffer for 5 min and subjected to SDS-PAGE analysis. Primary antibodies used for western blotting are listed in Supplementary Table 2.

CAGA-luc or SBE-luc transcriptional reporter assays

HEK293T, HeLa or HepG2 cells were seeded in the wells of 24-well plate (Corning) at a density of 3×10^5 and were transfected with 100 ng SMAD3/4-driven transcriptional CAGA-luc¹⁵ or BMP/TGF- β /SMAD-responsive SBE4-luc reporter⁵³, 80 ng β -galactosidase encoding plasmids and 320 ng indicated constructs using polyethyleneimine (PEI). For co-transfection of siRNAs and plasmids, HEK293T were seeded in the wells of 24-well plate (Corning) at a density of 3×10^5 and transfected with 20 ng SMAD3/4-driven transcriptional CAGA-luc reporter, 50 ng β -galactosidase encoding plasmids, 430 ng expression construct encoding OVOL1 or empty vector control, and non-targeting siRNA or siRNA SMARTpool targeting *SMAD7* (10 nM) using the DharmaFECT 1 transfection reagent (Horizon; Cat. Nr.: T-2001). At 16 h after transfection, cells were serum starved for 8 h and then kept in the presence or absence of TGF- β 3 (1 ng/ml) or BMP6 (10 ng/ml) overnight. Luciferase activity was measured using D-luciferin (Promega) as a substrate with a luminometer (PerkinElmer). The relative luciferase reporter activity was normalized to the β -galactosidase activity. All the experiments were repeated at least three times in biologically independent experiments, and representative results are shown.

MTS cell proliferation assays

MTS assay was carried out to evaluate the cell viability followed with manufacturer's instructions (Promega; Cat. Nr.: G3581). Cells were seed at a density of 1×10^3 cells in wells of 96-well plates (Corning). The absorbance of the samples was measured at 490 nm with a luminometer at 1, 2, 3, 4 and 5 days after seeding. Six biological replicates were included in each group.

Immunofluorescence staining

HeLa cells were transfected with constructs encoding T β RI-FLAG, OVOL1 or EGFP-SMAD7. Two days post-transfection, 4% paraformaldehyde (PFA) was used for fixing cells on covering glass for 20 min at RT, after which cells were permeabilized by phosphate buffered saline (PBS) supplemented with 0.1% Triton-X for 10 min. Afterwards, non-specific binding was blocked with 3% bovine serum albumin (BSA) dissolved in PBS for 1 h at RT. For staining of filamentous (F)-actin, cells were incubated with the Alexa Fluor 488 Phalloidin (Thermo Fisher Scientific; Cat. Nr.: A12379) in 1:500 dilution for 30 min at RT. To determine the localization of T β RI, OVOL1 and SMAD7, cells were incubated with primary antibodies in 1:100 dilution for 1 h at RT. After three times of washing with PBS, the specimens were probed with secondary antibodies (Invitrogen; Cat. Nr.: A21428 and A28175) in a dilution of 1:1000 for 1 h at RT. The specimens were then subjected to three washes with PBS and mounted with VECTASHIELD antifade mounting medium with DAPI (Vector Laboratories; Cat. Nr.: H-1200). The images were captured using a Leica SP8 confocal microscope (Leica Microsystems) and analyzed with the aid of LAS X software.

Ubiquitination assays

HEK293T cells transfected with indicated constructs or MDA-MB-231+Tet-ON OVOL1 cells and MCF10A-M2+Tet-ON shOVOL1 #1 cells stimulated with Dox for 2 days were treated 5 h prior to harvesting with 5 μ M proteasome inhibitor MG132 (Sigma-Aldrich; Cat. Nr.: 474787). After washing with cold PBS twice containing 10 mM N-ethylmaleimide (NEM; Sigma-Aldrich; Cat. Nr.: E3876), cells were lysed in 1% SDS-RIPA buffer (25 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% NP40, 0.5% Sodium Deoxycholate and 1% SDS) consisting of protease inhibitors and 10 mM NEM. Lysates were subsequently boiled for 5 min and diluted to 0.1% SDS at final concentration in RIPA buffer. Then protein concentrations were measured and the same amount of proteins was incubated with anti-FLAG agarose beads for 30 min at 4°C. For checking the ubiquitination of endogenous SMAD7, 1 μ L antibodies against SMAD7 were added to cell lysates and the mixtures were kept overnight at 4°C with rotation. The next day, 20 μ L Pierce™ Protein G Agarose (Thermo Fisher; Cat. Nr.: 20397) was added and kept for 2 h at 4°C with rotation. After five times of washing with RIPA buffer, beads were boiled in 2 \times loading buffer for 5 min and separated by SDS-PAGE. All the experiments were repeated at least three times in biologically independent experiments, and representative results are shown.

IncuCyte wound healing migration assays

Cells were seeded at a density of 5×10^4 in the wells of 96-well Essen ImageLock plate (Essen BioScience; Cat. Nr.: 4379). After 16 h, cells were serum starved with DMEM medium supplemented with 0.5% FBS for 8 h. The scratch wounds were generated by the Wound Maker (Essen BioScience). Floating cells were discarded by washing with PBS and attached cells were incubated with DMEM medium supplemented with 0.5% FBS. Migration was monitored in the IncuCyte live cell imaging system (Essen BioScience). Relative wound density was analyzed by the IncuCyte cell migration software. All the experiments were repeated at least three times in biologically independent experiments, and representative results are shown as mean \pm SD.

IncuCyte chemotaxis cell migration assay

In brief, 40 μ L MCF10A-M2 or SUM149PT cells suspended in medium supplemented with 0.5% serum were seeded at a density of 1×10^3 in the inserts of an IncuCyte clearview 96-well plate (Essen BioScience; Cat. Nr.: 4582). 20 μ L medium supplemented with 0.5% serum containing indicated compounds or corresponding vehicle controls was added into the inserts. Afterwards, cells were allowed to settle at ambient temperature for 20 min. In parallel, 200 μ L medium supplemented with 5% serum was added to the reservoir plates. Then the inserts containing cells were placed into a pre-filled plate. Cells on the top and bottom of inserts were imaged and analyzed using the IncuCyte system. The experiments were repeated for two times, and representative results are shown as mean \pm SD.

Immunohistochemical (IHC) staining and evaluation

Tissue microarray slides consisting of cancer adjacent breast tissues and matched breast carcinoma tissues (Biomax; Cat. Nr.: BR804b), and breast invasive carcinoma tissues with different grades (Biomax; Cat. Nr.: BC081116d) were dried overnight at 37°C. The next day, paraffine was removed by placing slides in xylene thrice, followed by placing in 100% ethanol twice. Endogenous peroxidase activity was blocked by 0.3% hydrogen peroxide for 20 min. Tissue sections were rehydrated in 96%, 70% and 50% Ethanol, respectively. Subsequently, antigen retrieval (10 min boiling in 0.01 M Sodium Citrate, pH 6.0) was performed after washing slides with PBS with 0.1% Tween 20 (PBST) for 5 min. Slides were cooled down to RT, followed by 3 times washing with PBST. Primary antibody against

OVOL1 (Thermo Fisher; Cat. Nr.: PA5-41480) diluted (1:200) in 1% BSA (dissolved in PBST) was applied to incubate slides overnight at 4°C. Slides were washed with PBST for 3 times before incubating with 1: 200 diluted biotinylated secondary antibody (DAKO; Cat. Nr.: E0353) for 30 min at RT. After 3 times washing with PBST, slides were subjected to Vectastain complex (Vector Laboratories; Cat. Nr.: PK-6100) incubation for 30 min at RT. Afterwards, slides were washed with PBST thrice and developed by DAB. Next, slides were counterstained with Mayers Haematoxylin (Sigma-Aldrich; Cat. Nr.: MHS80) for 10 s and dehydrated. Finally, Entellan (Merck; Cat. Nr.: 107961) was applied to mount the slides. Images were captured by slide scanner (3D Histech Panoramic 250). Staining was quantified and expressed as a H score which was determined by the formula $3 \times$ the percentage of strongly staining cells + $2 \times$ the percentage of moderately staining cells + the percentage of weakly staining cells.

Zebrafish extravasation assay of human breast cancer cells

Transgenic zebrafish lines Tg (fli1: EGFP) were raised according to standard procedures in compliance with the local Institutional Committee for Animal Welfare of the Leiden University. Zebrafish extravasation assays were performed as previously described⁴³. Zebrafish were washed with PBS twice and fixed with 4% PFA at 5 days after injection. Imaging of the zebrafish were carried out with the aid of an inverted SP5 STED confocal microscope (Leica). At least thirty zebrafish were analyzed for each group and representative images were taken. All the experiments were repeated for 2 times, and representative results are shown.

Mice xenograft model

22 five week-old female BALB/c nu/nu mice were ordered and acclimatized for one week in the Laboratory Animal Center (LAC) of the Netherlands Cancer Institute (Amsterdam, The Netherlands). All the mice were anesthetized by the inhalation of isoflurane (0.8 L/min) and intracardially injected with MDA-MB-231 luc⁺ Tet-ON OVOL1 cells (3×10^5 /100 μ l PBS) through the left heart ventricle. Three mice died after one day post-injection and the rest 19 mice were subdivided into two groups (9 in the -Dox groups and 10 in the +Dox group). Mice in the +Dox group were fed with 1% sucrose water supplemented with 2 mg/ml Doxycycline (Sigma; Cat. Nr.: D9891) to induce OVOL1 expression in MDA-MB-231 cells, while mice in the -Dox group were fed with the 1% sucrose water as a vehicle control. Bioluminescence imaging was carried out once a week to monitor the growth of metastases. All the mice experiment procedures were approved by the Animal Welfare Committee of the Netherlands Cancer Institute (Amsterdam, The Netherlands).

Statistical analyses

Statistical analysis was performed using Graphpad Prism 7 software. Results were expressed as the mean \pm SD of triplicates. All measurements were taken from distinct samples. For analysis, unpaired Student *t*-test was used and $p < 0.05$ was considered to be statistically significant. Paired Student *t*-test was carried out for analyzing the statistical significance of matched tissue samples in **Fig. 1d**. Two-way analysis of variance (ANOVA) was applied for statistical analysis of real time migration and chemotaxis assays. Pearsons' coefficient tests were performed to assess statistical significance for correlations between the expression of two genes or gene signatures. All measurements in this study were taken from distinct samples.

Data availability

The RNA-seq data presented in the study are deposited in GEO repository under the accession code GSE192548.

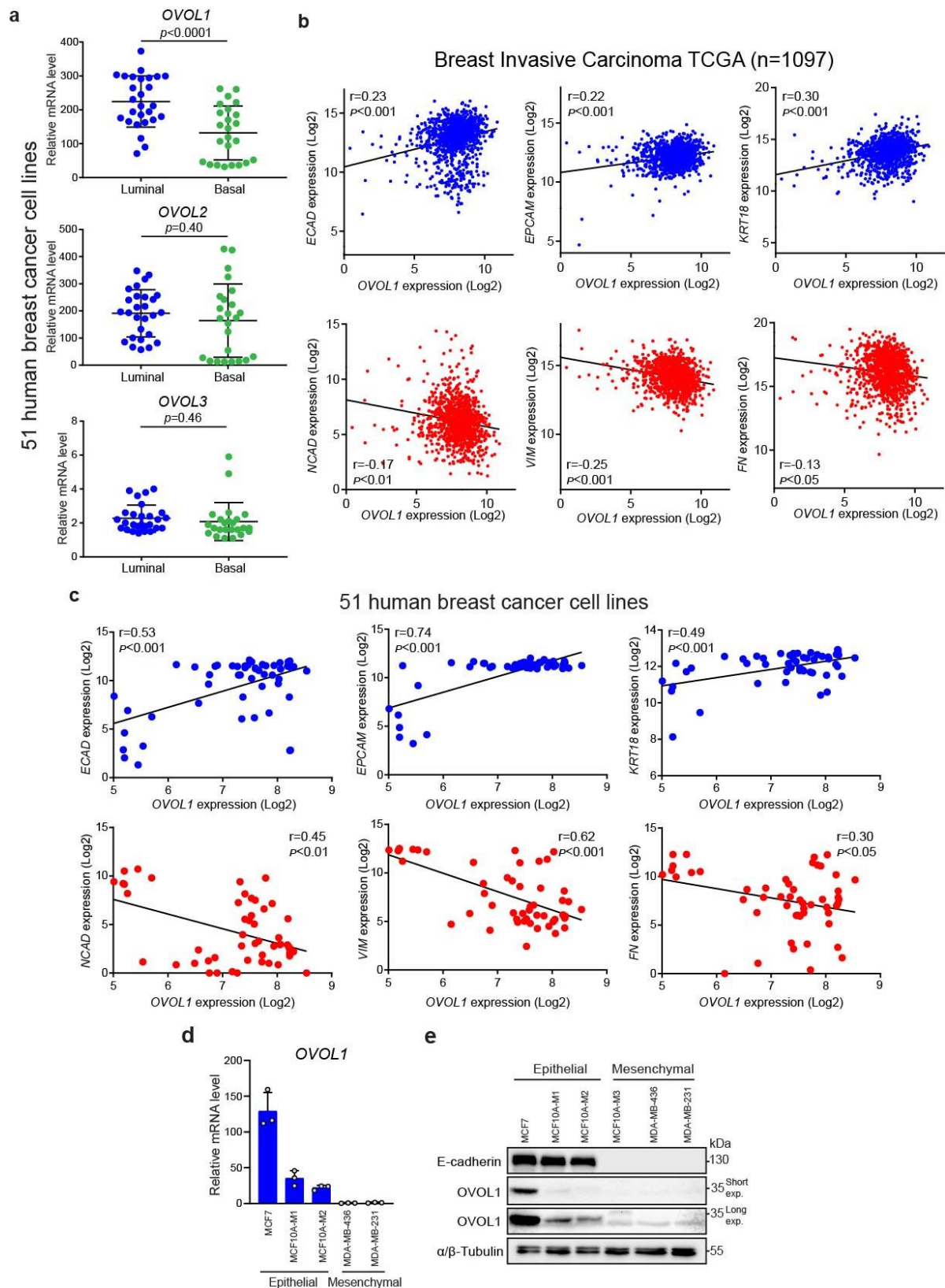
References

1. Siegel, R.L. *et al.* Colorectal cancer statistics, 2017. *CA Cancer J. Clin.* **67**, 177-193 (2017).
2. Wang, R. *et al.* The Clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer. *BMC cancer* **19**, 1091-1103 (2019).
3. Pastushenko, I. & Blanpain, C. EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol.* **29**, 212-226 (2019).
4. Dongre, A. & Weinberg, R.A. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* **20**, 69-84 (2019).
5. Hao, Y., Baker, D. & ten Dijke, P. TGF-beta-mediated epithelial-mesenchymal transition and cancer metastasis. *Int. J. Mol. Sci.* **20**, 2767-2801 (2019).
6. Shibue, T. & Weinberg, R.A. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* **14**, 611-629 (2017).
7. van Staalduinen, J., Baker, D., ten Dijke, P. & van Dam, H. Epithelial-mesenchymal-transition-inducing transcription factors: new targets for tackling chemoresistance in cancer? *Oncogene* **37**, 6195-6211 (2018).
8. Bolos, V. *et al.* The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J. Cell Sci.* **116**, 499-511 (2003).
9. Cano, A. *et al.* The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2**, 76-83 (2000).
10. Comijn, J. *et al.* The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol. Cell* **7**, 1267-1278 (2001).
11. Sha, Y.T. *et al.* Intermediate cell states in epithelial-to-mesenchymal transition. *Phys. Biol.* **16** (2019).
12. Yang, J. *et al.* Guidelines and definitions for research on epithelial-mesenchymal transition. *Nature reviews. Mol. Cell Biol.* **21**, 341-352 (2020).
13. Batlle, E. & Massague, J. Transforming growth factor-beta signaling in immunity and cancer. *Immunity* **50**, 924-940 (2019).
14. Grotendorst, G.R., Okochi, H. & Hayashi, N. A novel transforming growth factor beta response element controls the expression of the connective tissue growth factor gene. *Cell Growth Differ.* **7**, 469-480 (1996).
15. Denkler, S. *et al.* Direct binding of SMAD3 and SMAD4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. *EMBO J.* **17**, 3091-3100 (1998).
16. Gomez-Puerto, M.C., Iyengar, P.V., Garcia de Vinuesa, A., ten Dijke, P. & Sanchez-Duffhues, G. Bone morphogenetic protein receptor signal transduction in human disease. *J. Pathol.* **247**, 9-20 (2019).
17. Ren, J. *et al.* Reactivation of BMP signaling by suboptimal concentrations of MEK inhibitor and FK506 reduces organ-specific breast cancer metastasis. *Cancer lett.* **493**, 41-54 (2020).
18. Candia, A.F. *et al.* Cellular interpretation of multiple TGF-beta signals: intracellular antagonism between activin/BVg1 and BMP-2/4 signaling mediated by SMADs. *Development* **124**, 4467-4480 (1997).
19. Ehnert, S. *et al.* Transforming growth factor beta(1) inhibits bone morphogenetic protein (BMP)-2 and BMP-7 signaling via upregulation of Ski-related novel protein N (SnoN): possible mechanism for the failure of BMP therapy? *BMC Med.* **10**, 1-11 (2012).
20. Gronroos, E. *et al.* Transforming growth factor beta inhibits bone morphogenetic protein-induced transcription through novel phosphorylated SMAD1/5-SMAD3 complexes. *Mol. Cell. Biol.* **32**, 2904-2916 (2012).
21. Yan, X., Xiong, X. & Chen, Y.G. Feedback regulation of TGF-beta signaling. *Acta Biochim. Biophys. Sin.* **50**, 37-50 (2017).
22. Kavsak, P. *et al.* SMAD7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation. *Mol. Cell* **6**, 1365-1375 (2000).
23. Ebisawa, T. *et al.* Smurf1 interacts with transforming growth factor-beta type I receptor through SMAD7 and induces receptor degradation. *J. Biol. Chem.* **276**, 12477-12480 (2001).
24. van Capelle, C.D., Spit, M. & ten Dijke, P. Current perspectives on inhibitory SMAD7 in health and disease. *Crit. Rev. Biochem. Mol.* **55**, 691-715 (2020).
25. Koinuma, D. *et al.* Arkadia amplifies TGF-beta superfamily signalling through degradation of SMAD7. *EMBO J.* **22**, 6458-6470 (2003).
26. Zhang, L. *et al.* RNF12 Controls embryonic stem cell fate and morphogenesis in zebrafish embryos by targeting SMAD7 for degradation. *Mol. Cell* **46**, 650-661 (2012).
27. Lui, S.K.L. *et al.* USP26 regulates TGF-beta signaling by deubiquitinating and stabilizing SMAD7. *EMBO Rep.* **18**, 797-808 (2017).
28. Zhang, Z.K. *et al.* Breast cancer metastasis suppressor OTUD1 deubiquitinates SMAD7. *Nat. Commun.* **8**, 1-16 (2017).
29. Watanabe, K. *et al.* Mammary morphogenesis and regeneration require the inhibition of EMT at terminal end buds by Ovol2 transcriptional repressor. *Dev. Cell* **29**, 59-74 (2014).

30. Lee, B. *et al.* Transcriptional mechanisms link epithelial plasticity to adhesion and differentiation of epidermal progenitor cells. *Dev. Cell* **29**, 47-58 (2014).
31. Li, B.A. *et al.* Ovoll regulates meiotic pachytene progression during spermatogenesis by repressing Id2 expression. *Development* **132**, 1463-1473 (2005).
32. Saxena, K., Srikrishnan, S., Celia-Terrassa, T. & Jolly, M.K. OVOL1/2: Drivers of Epithelial Differentiation in Development, Disease, and Reprogramming. *Cells Tissues Organs* 1-10 (2020).
33. Nair, M., Bilanchone, V., Ortt, K., Sinha, S. & Dai, X. Ovoll represses its own transcription by competing with transcription activator c-Myb and by recruiting histone deacetylase activity. *Nucleic Acids Res.* **35**, 1687-1697 (2007).
34. Teng, A., Nair, M., Wells, J., Segre, J.A. & Dai, X. Strain-dependent perinatal lethality of Ovoll-deficient mice and identification of Ovoll2 as a downstream target of Ovoll1 in skin epidermis. *Biochim. Biophys. Acta* **1772**, 89-95 (2007).
35. Roca, H. *et al.* Transcription factors OVOL1 and OVOL2 induce the mesenchymal to epithelial transition in human cancer. *PLoS One* **8**, 1-20 (2013).
36. Wu, R. *et al.* OVOL2 antagonizes TGF- β signaling to regulate epithelial to mesenchymal transition during mammary tumor metastasis. *Oncotarget* **8**, 39401-39416 (2017).
37. Neve, R.M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515-527 (2006).
38. Irshad, S. *et al.* Bone morphogenetic protein and Notch signalling crosstalk in poor-prognosis, mesenchymal-subtype colorectal cancer. *J. Pathol.* **242**, 178-192 (2017).
39. Koboldt, D.C. *et al.* Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012).
40. Gyorffy, B. *et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. and Treat.* **123**, 725-731 (2010).
41. Nagy, A., Munkacsy, G. & Gyorffy, B. Pancancer survival analysis of cancer hallmark genes. *Sci. Rep.* **11**, 6047 (2021).
42. Xie, B.B. *et al.* MOBCdb: a comprehensive database integrating multi-omics data on breast cancer for precision medicine. *Breast Cancer Res. and Treat.* **169**, 625-632 (2018).
43. Ren, J., Liu, S., Cui, C. & ten Dijke, P. Invasive behavior of human breast cancer cells in embryonic zebrafish. *J. Vis. Exp.* **122**, 1-12 (2017).
44. Kowanz, M., Valcourt, U., Bergstrom, R., Heldin, C.R.F. & Moustakas, A. Id2 and Id3 define the potency of cell proliferation and differentiation responses to transforming growth factor beta and bone morphogenetic protein. *Mol. Cell. Biol.* **24**, 4241-4254 (2004).
45. Gomis, R.R. *et al.* A FoxO-SMAD synexpression group in human keratinocytes. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12747-12752 (2006).
46. Descargues, P. *et al.* IKK alpha is a critical coregulator of a SMAD4-independent TGF beta-SMAD2/3 signaling pathway that controls keratinocyte differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 2487-2492 (2008).
47. Cuny, G.D. *et al.* Structure-activity relationship study of bone morphogenetic protein (BMP) signaling inhibitors. *Bioorg. Med. Chem. Lett.* **18**, 4388-4392 (2008).
48. Laping, N.J. *et al.* Inhibition of transforming growth factor (TGF)-beta 1-induced extracellular matrix with a novel inhibitor of the TGF-beta type I receptor kinase activity: SB-431542. *Mol. Pharmacol.* **62**, 58-64 (2002).
49. Lagna, G., Hata, A., HemmatiBrivanlou, A. & Massague, J. Partnership between DPC4 and SMAD proteins in TGF-beta signalling pathways. *Nature* **383**, 832-836 (1996).
50. Zhang, Y., Musci, T. & Derynck, R. The tumor suppressor SMAD4 DPC 4 as a central mediator of SMAD function. *Curr. Biol.* **7**, 270-276 (1997).
51. Takase, M. *et al.* Induction of SMAD6 mRNA by bone morphogenetic proteins. *Biochem. Biophys. Res. Commun.* **244**, 26-29 (1998).
52. Nakao, A. *et al.* Identification of SMAD7, a TGFbeta-inducible antagonist of TGF-beta signalling. *Nature* **389**, 631-635 (1997).
53. Jonk, L.J., Itoh, S., Heldin, C.H., ten Dijke, P. & Kruijer, W. Identification and functional characterization of a SMAD binding element (SBE) in the JunB promoter that acts as a transforming growth factor-beta, activin, and bone morphogenetic protein-inducible enhancer. *J. Biol. Chem.* **273**, 21145-21152 (1998).
54. Padua, D. *et al.* TGF beta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**, 66-77 (2008).
55. Komander, D. & Rape, M. The Ubiquitin Code. *Annu. Rev. Biochem.* **81**, 203-229 (2012).
56. Tsuji, G. *et al.* Aryl hydrocarbon receptor activation restores filaggrin expression via OVOL1 in atopic dermatitis. *Cell Death Dis.* **8**, 1-8 (2017).

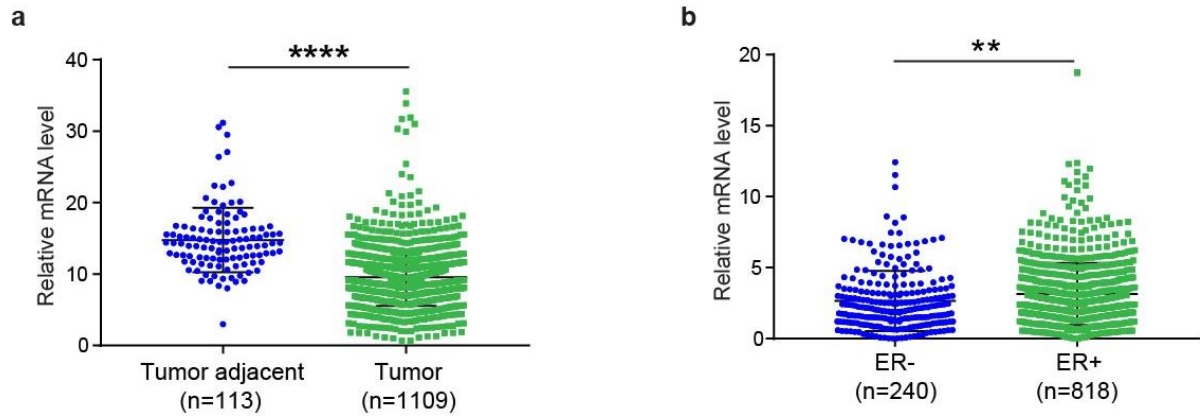
57. Jia, D.Y. *et al.* OVOL guides the epithelial-hybrid-mesenchymal transition. *Oncotarget* **6**, 15436-15448 (2015).
58. Jolly, M.K. *et al.* Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget* **7**, 27067-27084 (2016).
59. Murata, M. *et al.* OVOL2-mediated ZEB1 downregulation may prevent promotion of actinic keratosis to cutaneous squamous cell carcinoma. *J. Clin. Med.* **9**, 1-16 (2020).
60. Xu, C.F., Yan, T.Y. & Yang, J.R. OVOL1 inhibits oral squamous cell carcinoma growth and metastasis by suppressing zinc finger E-box binding homeobox 1. *Int. J. Clin. and Exp. Pathol.* **12**, 2801-2808 (2019).
61. Pattabiraman, D.R. *et al.* Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. *Science* **351**, 1042-+ (2016).
62. Lombaerts, M. *et al.* E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. *Br. J. Cancer* **94**, 661-671 (2006).
63. Ye, G.D. *et al.* OVOL2, an Inhibitor of WNT Signaling, Reduces Invasive Activities of Human and Mouse Cancer Cells and Is Down-regulated in Human Colorectal Tumors. *Gastroenterology* **150**, 659-671 (2016).
64. Li, B.A. *et al.* The LEF1/beta-catenin complex activates movo1, a mouse homolog of *Drosophila ovo* required for epidermal appendage differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 6064-6069 (2002).
65. Min, H.Y., Sung, Y.K., Kim, E.J. & Jang, W.G. OVO homologue-like 1 promotes osteoblast differentiation through BMP2 expression. *J. Cell. Physiol.* **234**, 11842-11849 (2019).
66. Renaud, S.J. *et al.* OVO-like 1 regulates progenitor cell fate in human trophoblast development. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6175-6184 (2015).
67. Rembold, M. *et al.* A conserved role for Snail as a potentiator of active transcription. *Genes Dev.* **28**, 167-181 (2014).
68. Massague, J. TGF beta signalling in context. *Nat. Rev. Mol. Cell Biol.* **13**, 616-630 (2012).
69. Kiyomatsu-Oda, M., Uchi, H., Morino-Koga, S. & Furue, M. Protective role of 6-formylindolo[3,2-b]carbazole (FICZ), an endogenous ligand for arylhydrocarbon receptor, in chronic mite-induced dermatitis. *J Dermatol Sci* **90**, 284-294 (2018).
70. Schulz, V.J. *et al.* Non-dioxin-like AhR ligands in a mouse peanut allergy model. *Toxicol Sci* **128**, 92-102 (2012).
71. Wheeler, J.L., Martin, K.C., Resseguie, E. & Lawrence, B.P. Differential consequences of two distinct AhR ligands on innate and adaptive immune responses to influenza A virus. *Toxicol Sci* **137**, 324-334 (2014).
72. Cheng, S.H.C. *et al.* Validation of the 18-gene classifier as a prognostic biomarker of distant metastasis in breast cancer. *Plos One* **12**, 1-15 (2017).
73. Caldon, C.E. *et al.* Cyclin E2 Overexpression Is Associated with Endocrine Resistance but not Insensitivity to CDK2 Inhibition in Human Breast Cancer Cells. *Mol. Cancer Ther.* **11**, 1488-1499 (2012).
74. Bos, P.D. *et al.* Genes that mediate breast cancer metastasis to the brain. *Nature* **459**, 1005-U1137 (2009).
75. Miller, L.D. *et al.* An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17882-17882 (2005).
76. Hoeflich, K.P. *et al.* In vivo Antitumor Activity of MEK and Phosphatidylinositol 3-Kinase Inhibitors in Basal-Like Breast Cancer Models. *Clin. Cancer Res.* **15**, 4649-4664 (2009).
77. Liu, S. *et al.* Deubiquitinase Activity Profiling Identifies UCHL1 as a Candidate Oncoprotein That Promotes TGFbeta-Induced Breast Cancer Metastasis. *Clin. Cancer Res.* **26**, 1460-1473 (2020).
78. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545-15550 (2005).
79. Mootha, V.K. *et al.* PGC-1 alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267-273 (2003).
80. Padua, D. *et al.* TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**, 66-77 (2008).
81. Pangas, S.A. *et al.* Conditional deletion of SMAD1 and SMAD5 in somatic cells of male and female gonads leads to metastatic tumor development in mice. *Mol. Cell Biol.* **28**, 248-257 (2008).

Supplementary information

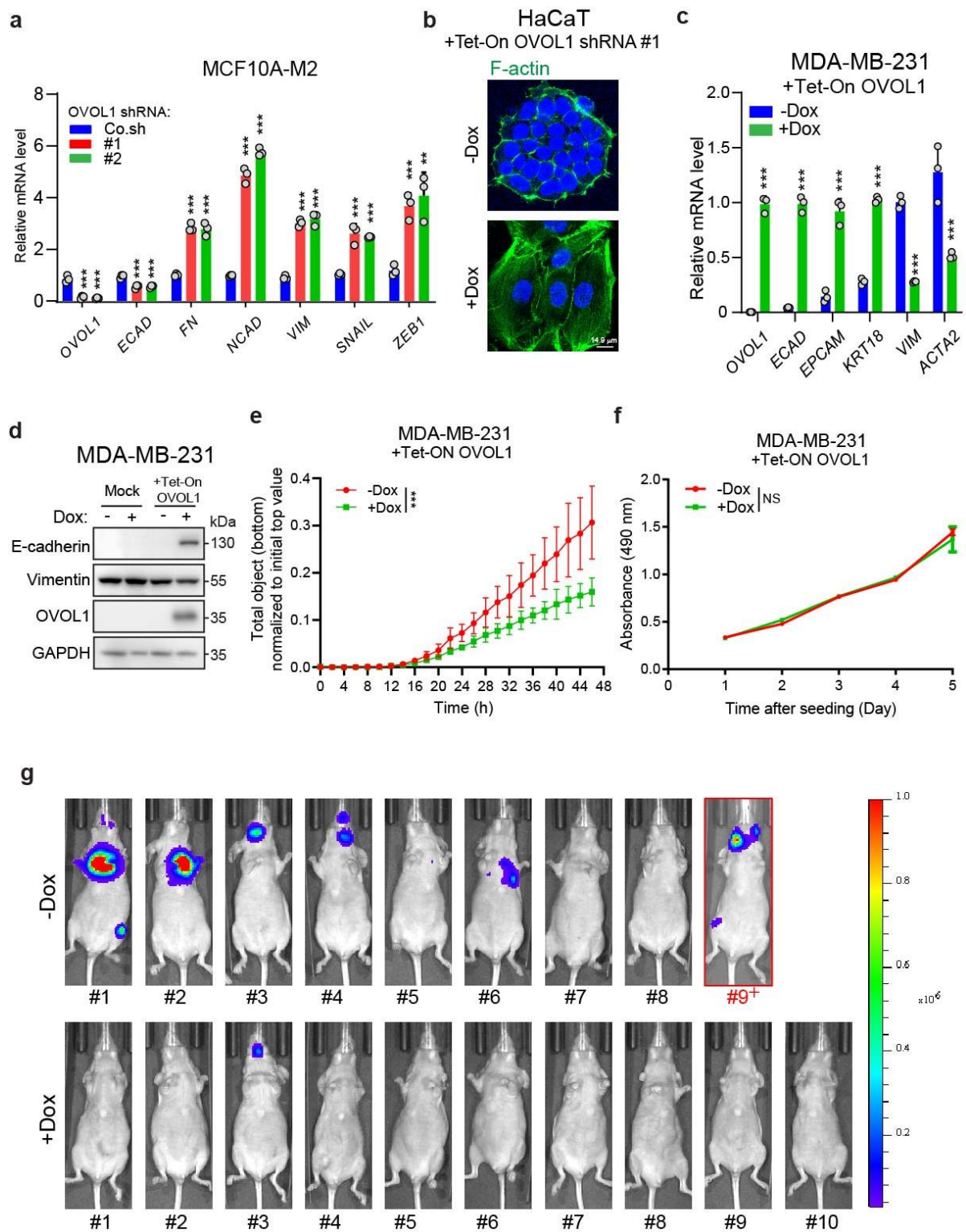


Supplementary Fig. 1 *OVOL1* expression is positively correlated with the expression of epithelial markers while negatively correlated with the expression of mesenchymal markers. (a) *OVOL1*, *OVOL2* or *OVOL3* mRNA expression in 51 breast cell lines that were subdivided into luminal and basal groups. (b, c) Scatter plots demonstrating positive correlation between the expression of *OVOL1* and epithelial markers (*ECAD*, *EPCAM* or *KRT18*) or inverse correlation with mesenchymal

markers (*NCAD*, *VIM* or *FN*) in datasets consisting of TCGA breast invasive carcinoma samples (**b**) or 51 human breast cancer cell lines (**c**). (**d**) RT-qPCR analysis of *OVOL1* expression in normal breast cells (MCF10A-M1), pre-malignant breast cells (MCF10A-M2) and luminal breast cancer cells (MCF7) with an epithelial phenotype and triple negative breast cancer cell lines with an aggressive mesenchymal phenotype (MDA-MB-436 and MDA-MB-231). The result is presented as mean \pm SD in technical triplicates. (**e**) Western blot analysis of E-cadherin and *OVOL1* protein levels in normal breast cells (MCF10A-M1), pre-malignant breast cells (MCF10A-M2) and luminal breast cancer cells (MCF7) with an epithelial phenotype, and triple negative breast cancer cell lines with an aggressive mesenchymal phenotype (MDA-MB-436 and MDA-MB-231). α/β -tubulin levels were analyzed to control for equal loading.



Supplementary Fig. 2 *OVOL1* is lower expressed in tumor samples with poor prognosis. (**a**) Differential *OVOL1* mRNA expression in tumor adjacent normal breast tissue samples and tumor samples from breast cancer patients. (**b**) Differential analysis of *OVOL1* mRNA expression in breast cancer patient specimens that were subdivided into ER negative (ER-) and ER positive (ER+) groups. The results in (**a**) and (**b**) are expressed as mean \pm SD. ** $0.001 < p < 0.01$, **** $p < 0.0001$.

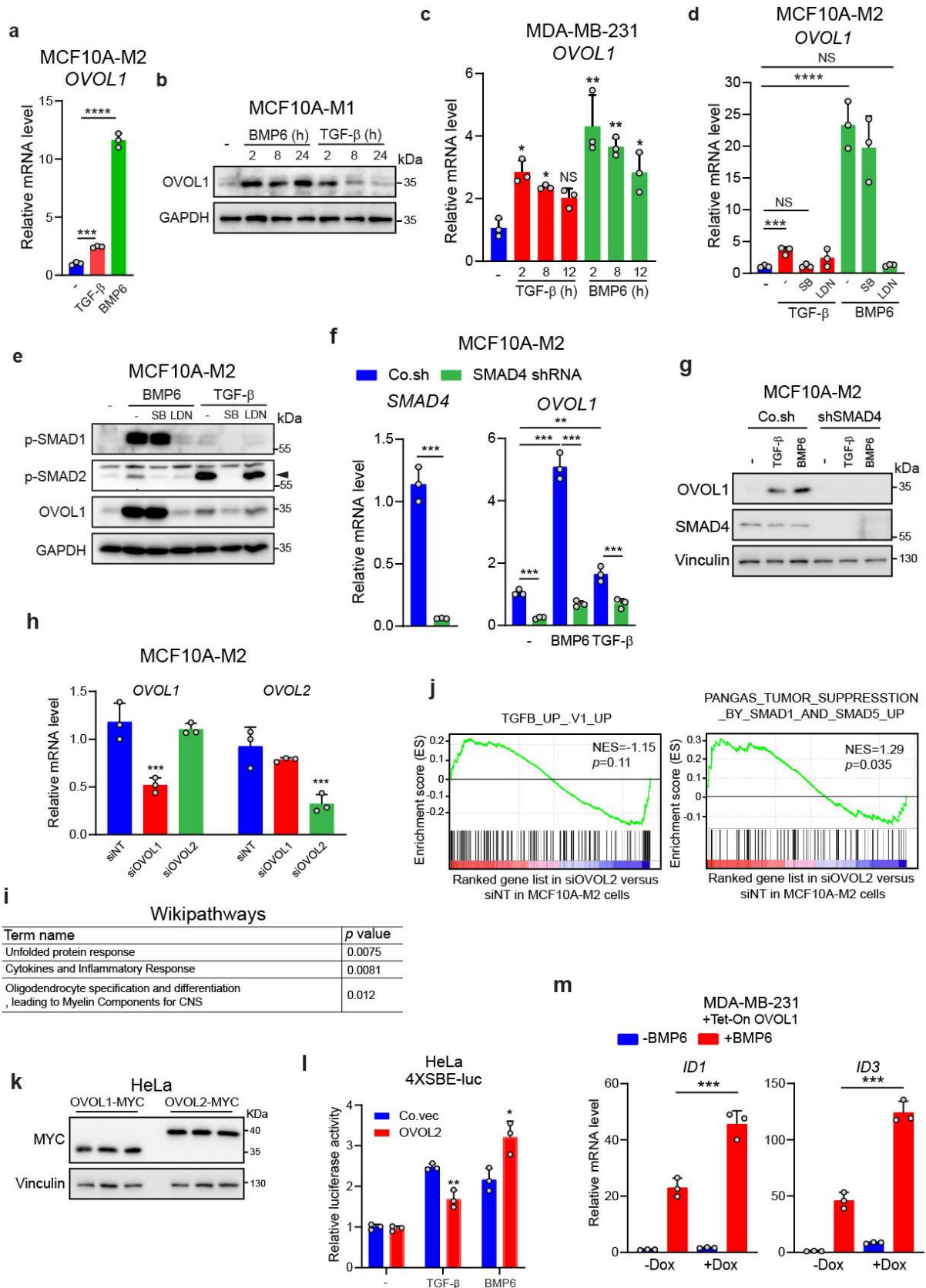


Supplementary Fig. 3 OVOL1 inhibits the EMT and MDA-MB-231 breast cancer cell migration and metastasis but not MDA-MB-231 cell viability. (a) Measurement of EMT markers expression by RT-qPCR in MCF10A-M2 cells upon OVOL1 knockdown. Statistical analyses were performed between groups of vector control (Co.sh) and OVOL1 sh#1 or sh#2, respectively. The results are expressed as mean \pm SD. ** $0.001 < p < 0.01$, *** $0.0001 < p < 0.001$. (b) Immunofluorescence detection of F-actin and 4, 6-diamidino-2-phenylindole (DAPI) staining of HaCaT cells upon OVOL1 knockdown induced by Doxycycline (Dox). Cells were treated without or with Dox for 48 h. Effectivity of *OVOL1* gene knockdown was confirmed (Supplementary Fig. 3a). Scale bar, 14.9 μ m. (c) RT-qPCR detection of EMT markers expression in MDA-MB-231 cells in the absence or presence of OVOL1 ectopic expression. Cells were either not stimulated or stimulated with Doxycycline (Dox) for 2 days. Statistical analyses were carried out between -Dox and +Dox groups. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$. (d) Analysis expression of EMT markers by western blotting in MDA-MB-231 cells without (Mock) or

Chapter 4

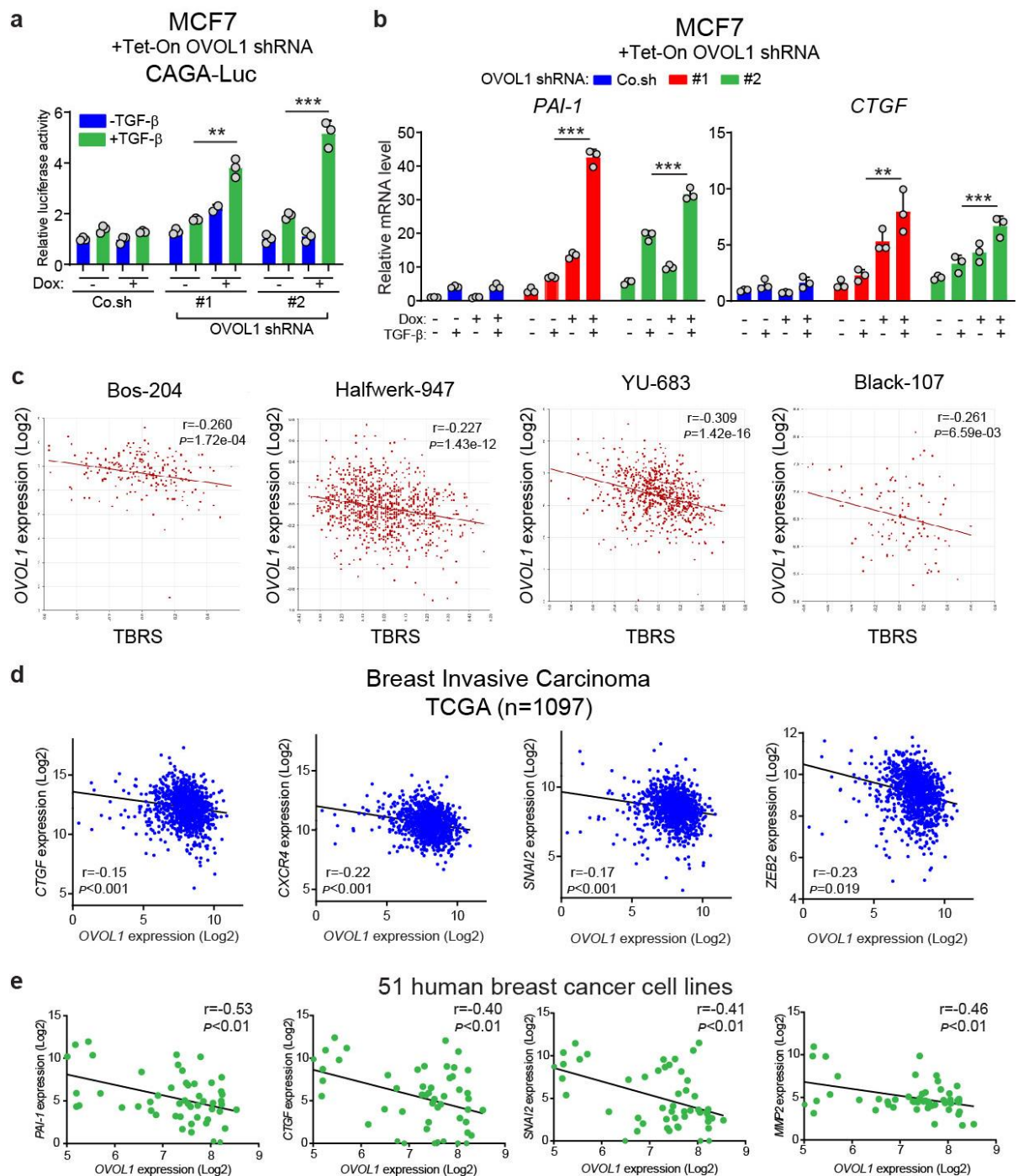
with inducible *OVOL1* expression (+Tet-On *OVOL1*). Cells were treated without or with Doxycycline (Dox) for 2 days. To control for equal loading GAPDH levels were analyzed. (e) IncuCyte chemotaxis assay to evaluate the migratory abilities of MDA-MB-231 cells upon the ectopic expression of *OVOL1*. Cells were stimulated without or with Doxycycline (Dox) two days before being seeded in the chambers. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$, NS, not significant. (f) MTS cell viability analysis of MDA-MB-231 cells without or with ectopic expression of *OVOL1*. Cells were kept in the presence or absence of Doxycycline (Dox) for indicated time points. See Supplementary Fig. 3c for the analysis of *OVOL1* expression after Dox challenge. (g) Whole body bioluminescence images (BLI) of mice treated without or with Dox. Images of all but one mice at 9 weeks are shown; one mouse without Dox treatment was imaged at 8 weeks and thereafter terminated. Three selected mice without and with Dox are shown in Fig. 2f.

OVOL1 inhibits breast cancer cell invasion by enhancing the degradation of TGF- β type I receptor

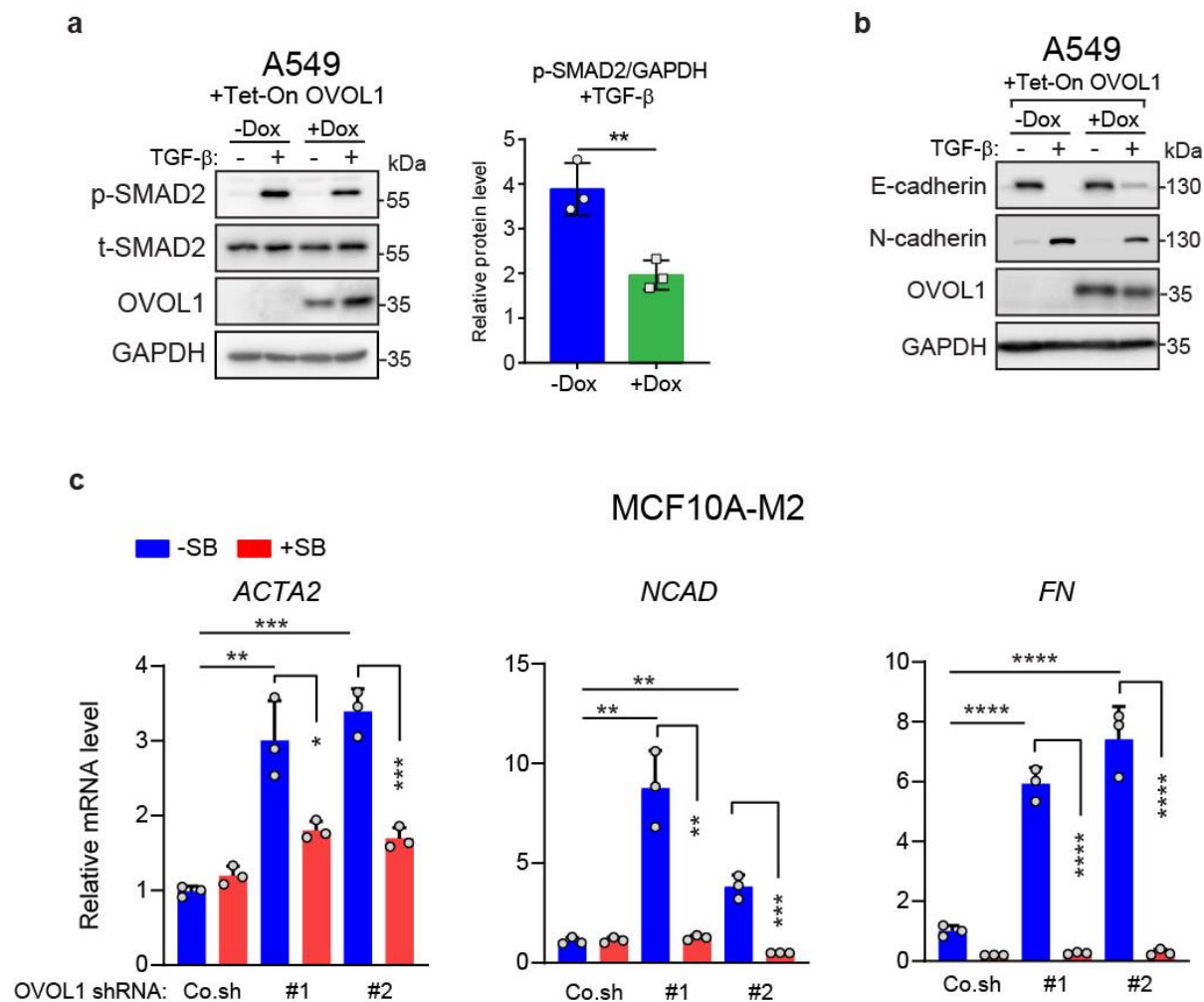


Supplementary Fig. 4 *OVOL1* is a target gene of BMP/SMAD and TGF- β /SMAD pathways. (a) *OVOL1* expression detected by RT-qPCR in MCF10A-M2 cells upon the stimulation of TGF- β (5 ng/ml) or BMP6 (50 ng/ml) for indicated time points. Statistical analyses were performed between control group and indicated groups with ligand treatment. The results are expressed as mean \pm SD. *** 0.0001 < p < 0.001, **** p < 0.0001. (b) Western blotting detection of *OVOL1* expression in MCF10A cells treated with vehicle control (-) or TGF- β (5 ng/ml) or BMP6 (50 ng/ml) for indicated time points. To control

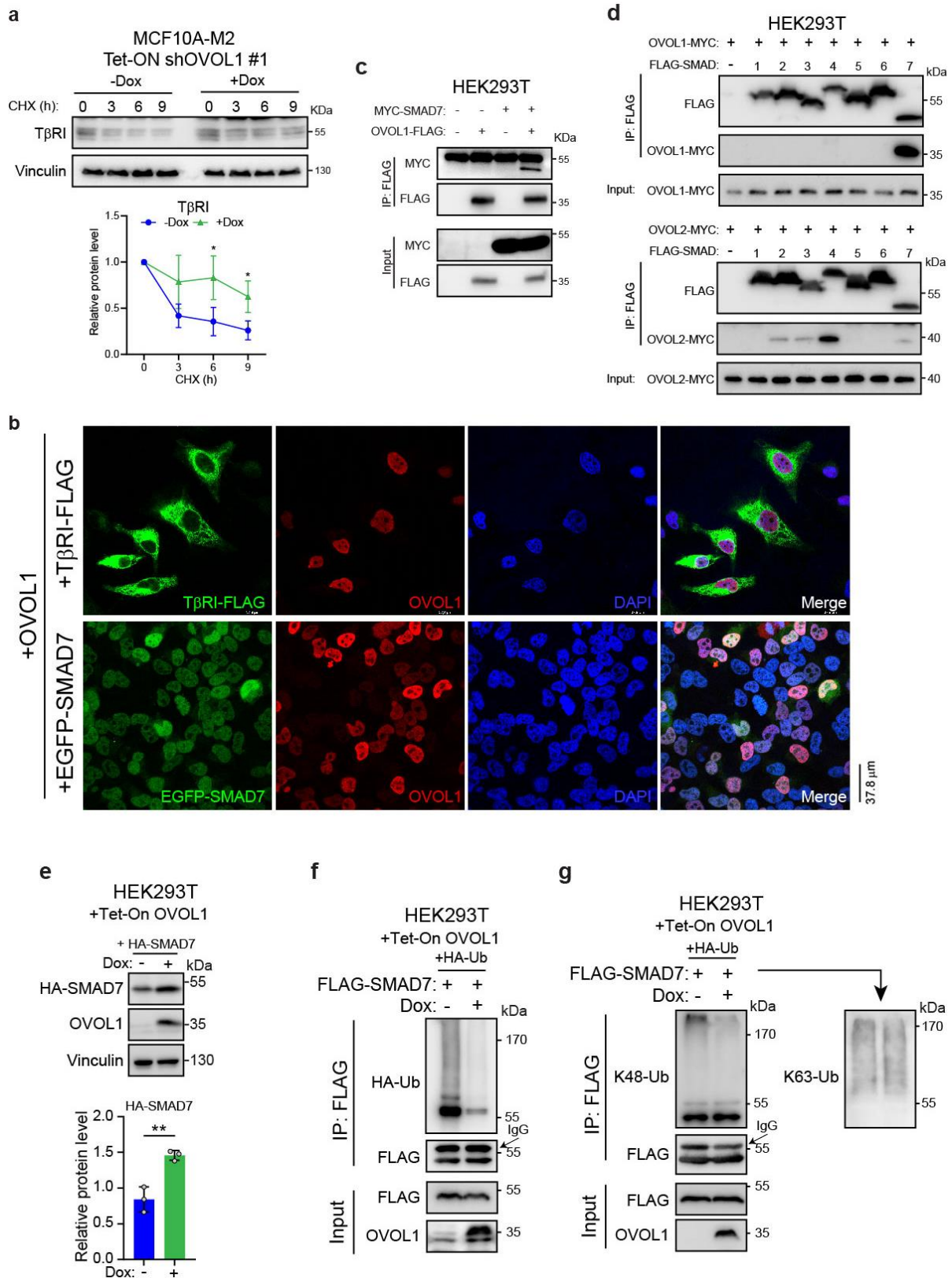
for equal loading GAPDH levels were analyzed. (c) *OVOL1* expression detected by RT-qPCR in MDA-MB-231 cells upon the stimulation of TGF- β (5 ng/ml) or BMP6 (50 ng/ml) for indicated time points. Statistical analyses were performed between control group and indicated groups with ligand treatment. The results are expressed as mean \pm SD. * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$, NS, not significant. (d) *OVOL1* expression quantified by RT-qPCR in MCF10A-M2 cells. Cell were either not treated or treated with small molecule kinase inhibitors of BMP type I receptor (LDN193189; LDN; 120 nM) or TGF- β type I receptor (SB431542; SB; 5 μ M) for 30 min followed by the stimulation of TGF- β (5 ng/ml) or BMP6 (50 ng/ml) for 2 h. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$, **** $p < 0.0001$, NS, not significant. (e) Western blot analysis of *OVOL1* protein expression in MCF10A-M2 cells. Cells were kept in the presence or absence of inhibitors of BMP (LDN193189; LDN) or TGF- β (SB431542; SB) for 30 min followed by the treatment of vehicle control (-), TGF- β (5 ng/ml) or BMP6 (50 ng/ml) for 4 h. The phosphorylation of SMAD1 (p-SMAD1) or SMAD2 (p-SMAD2) was detected to confirm the activation of the BMP or TGF- β pathway, respectively. To control for equal loading GAPDH levels were analyzed. (f) *OVOL1* expression quantification by RT-qPCR in MCF10A-M2 cells upon SMAD4 depletion. Cells were serum starved overnight and treated with vehicle control (-), TGF- β (5 ng/ml) or BMP6 (50 ng/ml) for 2 h. The results are expressed as mean \pm SD. ** $0.001 < p < 0.01$, *** $0.0001 < p < 0.001$. (g) Western blotting assay for detecting *OVOL1* expression in MCF10A-M2 cells without (Co.sh) or with (shSMAD4) SMAD4 depletion. Cells were serum starved overnight and treated with vehicle control (-), TGF- β (5 ng/ml) or BMP6 (50 ng/ml) for 4 h. To control for equal loading Vinculin levels were analyzed. (h) RT-qPCR evaluation of *OVOL1* or *OVOL2* levels in MCF10A-M2 cells transfected with the indicated siRNAs. Comparisons were performed against the siNT group. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$. (i) The pathway enrichment results from wikipathways when *OVOL2* was depleted in MCF10A-M2 cells. (j) GSEA analyses of the correlations between (manipulated) *OVOL2* expression level and the TGF- β (left) or SMAD1/5 (BMP; right) gene response signature. (k) Western blotting quantification for *OVOL1* and *OVOL2* expression in HeLa cells transfected with MYC-tagged *OVOL1* or *OVOL2*. Cell lysates in three biological replicates from the luciferase assays in Fig. 3e and Supplementary Fig. 4l were analyzed. To control for equal loading Vinculin levels were analyzed. (l) Quantification of the luciferase transcriptional activity in HeLa cells transfected with BMP/TGF- β -responsive SBE4-luc reporter and empty vector (Co.vec) or *OVOL2*. The results are expressed as mean \pm SD. * $0.01 < p < 0.05$, ** $0.001 < p < 0.001$. (m) Measurement of *ID1* and *ID3* expression by RT-qPCR in MDA-MB-231 cells with inducible *OVOL1* expression. Cells were either not treated or treated with Doxycycline (Dox) for 2 days before serum starvation overnight and adding BMP6 (50 ng/ml) for 2 h. The results are expressed as mean \pm SD. *** $0.0001 < p < 0.001$.



Supplementary Fig. 5 OVOL1 alleviates TGF- β pathway transduction. (a) Reporter assay for quantifying the luciferase activity in MCF7 cells stably expressing TGF- β -induced SMAD3/4-dependent CAGA-luc transcriptional reporter. Cells without (Co.sh) or with inducible OVOL1 knockdown (shRNA #1 and #2) were kept in the presence or absence of Doxycycline (Dox) for 2 days. Subsequently, cells were serum starved for 8 h and stimulated with vehicle control (-) or TGF- β (5 ng/ml) overnight. The results are expressed as mean \pm SD. ** 0.001 < p < 0.001, *** 0.0001 < p < 0.001. (b) RT-qPCR measurement of *PAI-1* and *CTGF* expression in MCF7 cells with empty vector control (Co.sh) or inducible OVOL1 knockdown (shRNA #1 and #2). Cells were treated without or with Doxycycline (Dox) for 2 days before serum starvation overnight and stimulated with vehicle control (-) or TGF- β (5 ng/ml) for 4 h. The results are expressed as mean \pm SD. ** 0.001 < p < 0.001, *** 0.0001 < p < 0.001. (c) Scatter plot illustrating the inverse correlation between *OVOL1* and the TGF- β response signature (TBRS) in four breast cancer datasets. Titles on top of each panels indicated the datasets in which the RNA-seq were analyzed. (d) Scatter plot of inverse correlation between *OVOL1* and the TGF- β /SMAD target genes (*CTGF*, *CXCR4*, *SNAI2* or *ZEB2*) in a dataset consisting of TCGA breast invasive carcinoma¹. (e) Scatter plot showing correlation between *OVOL1* and the TGF- β /SMAD target genes (*PAI-1*, *CTGF*, *SNAI2* or *MMP2*) in a dataset consisting of 51 human breast cancer cell lines.

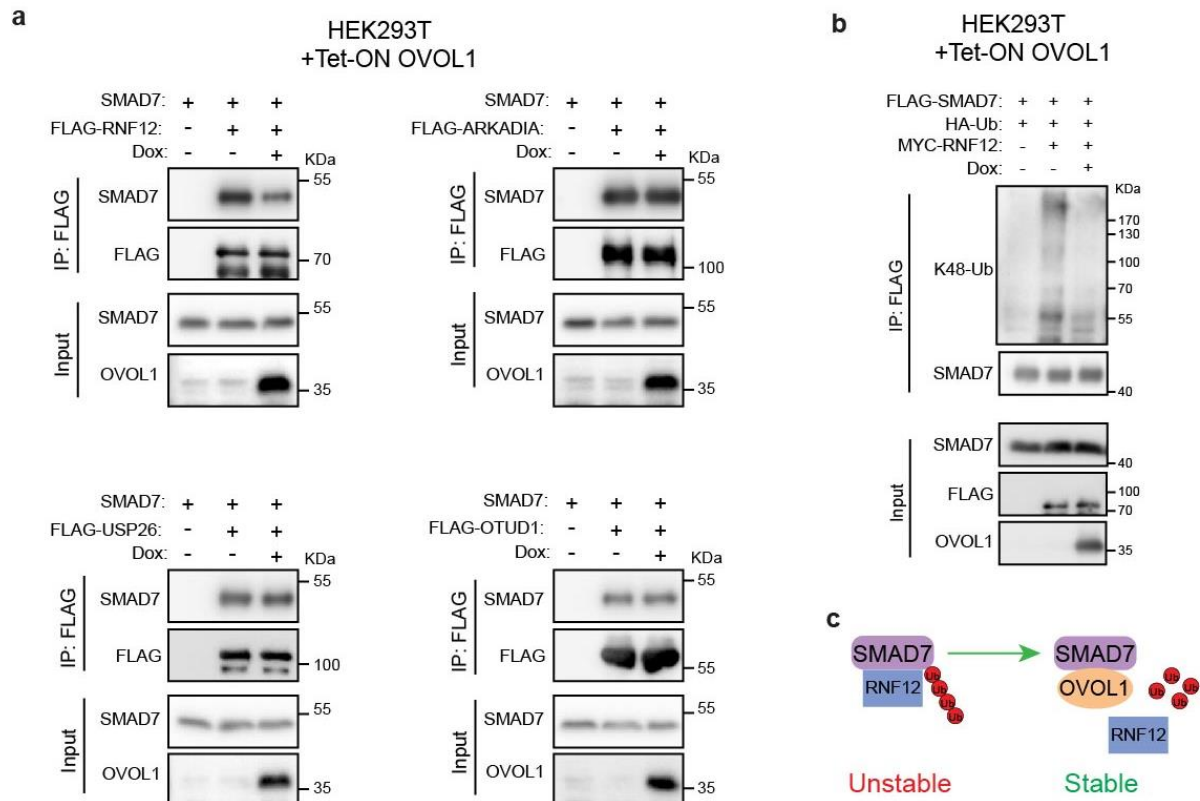


Supplementary Fig. 6 OVOL1 attenuates TGF-β signaling transduction. (a) Western blotting detection of the phosphorylation of SMAD2 (p-SMAD2) and total SMAD2 (t-SMAD2) in A549 cells without or with OVOL1 ectopic expression induced by Doxycycline (Dox). To control for equal loading GAPDH levels were analyzed (left panel). Cells were kept in the presence or absence of Dox for 2 days before serum starvation overnight and stimulation with vehicle control (-) or TGF-β (1 ng/ml) for 2 h. Relative protein level of p-SMAD2 was quantified from three independent sets of experiments (right panel). The results are expressed as mean ± SD. ** 0.001 < *p* < 0.001. (b) The expression of E-cadherin and N-cadherin detected by Western blotting in A549 cells without or with ectopic expression of OVOL1. Cells were either not treated or treated with Doxycycline (Dox) for 2 days before stimulation with vehicle control (-) or TGF-β (1 ng/ml) for 1 day. To control for equal loading GAPDH levels were analyzed. (c) RT-qPCR quantification of mesenchymal markers expression in MCF10A-M2 cells upon OVOL1 knockdown. Cells were kept in the presence or absence of SB431542 (SB) for 2 days. The results are expressed as mean ± SD. * 0.01 < *p* < 0.05, ** 0.001 < *p* < 0.001, *** 0.0001 < *p* < 0.001, **** *p* < 0.0001.

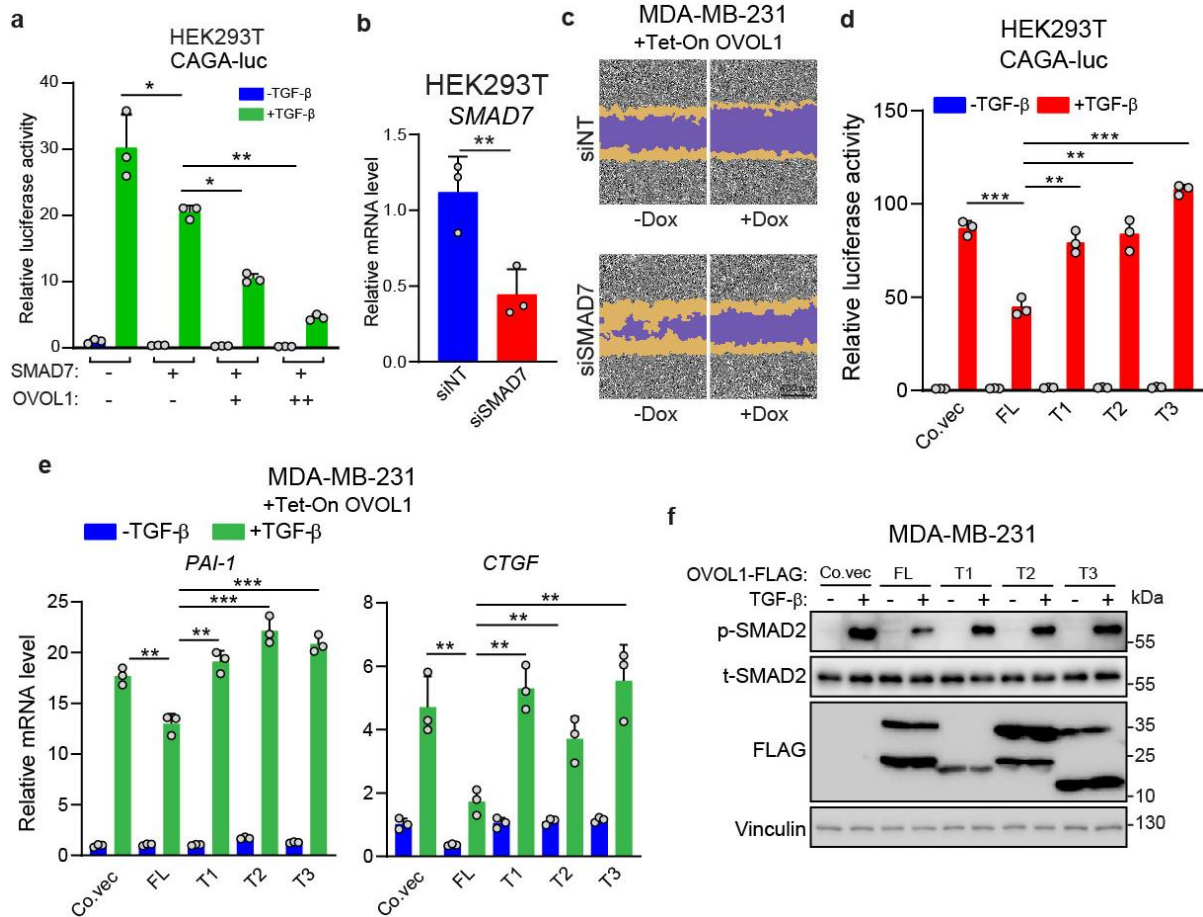


Supplementary Fig. 7 OVOL1 interacts with SMAD7. (a) Western blot analysis of T β RI expression levels in MCF10A-M2 cells upon OVOL1 knockdown induced by Doxycycline (Dox). Cells were either not treated or treated with Doxycycline (Dox) for 2 days followed by the stimulation of cycloheximide (CHX; 50 μ g/ml) for indicated time points. Quantification of the relative T β RI protein level is shown in the lower panel. Statistical analyses were performed at indicated time points. The results are expressed as mean \pm SD. * 0.01 < p < 0.05. To control for equal loading Vinculin levels were analyzed. (b) Immunofluorescence detection of T β RI-FLAG, OVOL1, EGFP-SMAD7 and 4, 6-diamidino-2-phenylindole (DAPI) staining of HeLa cells upon the transfection of indicated constructs. Scale bar, 37.8 μ m. (c) Western blot analysis of MYC-SMAD7

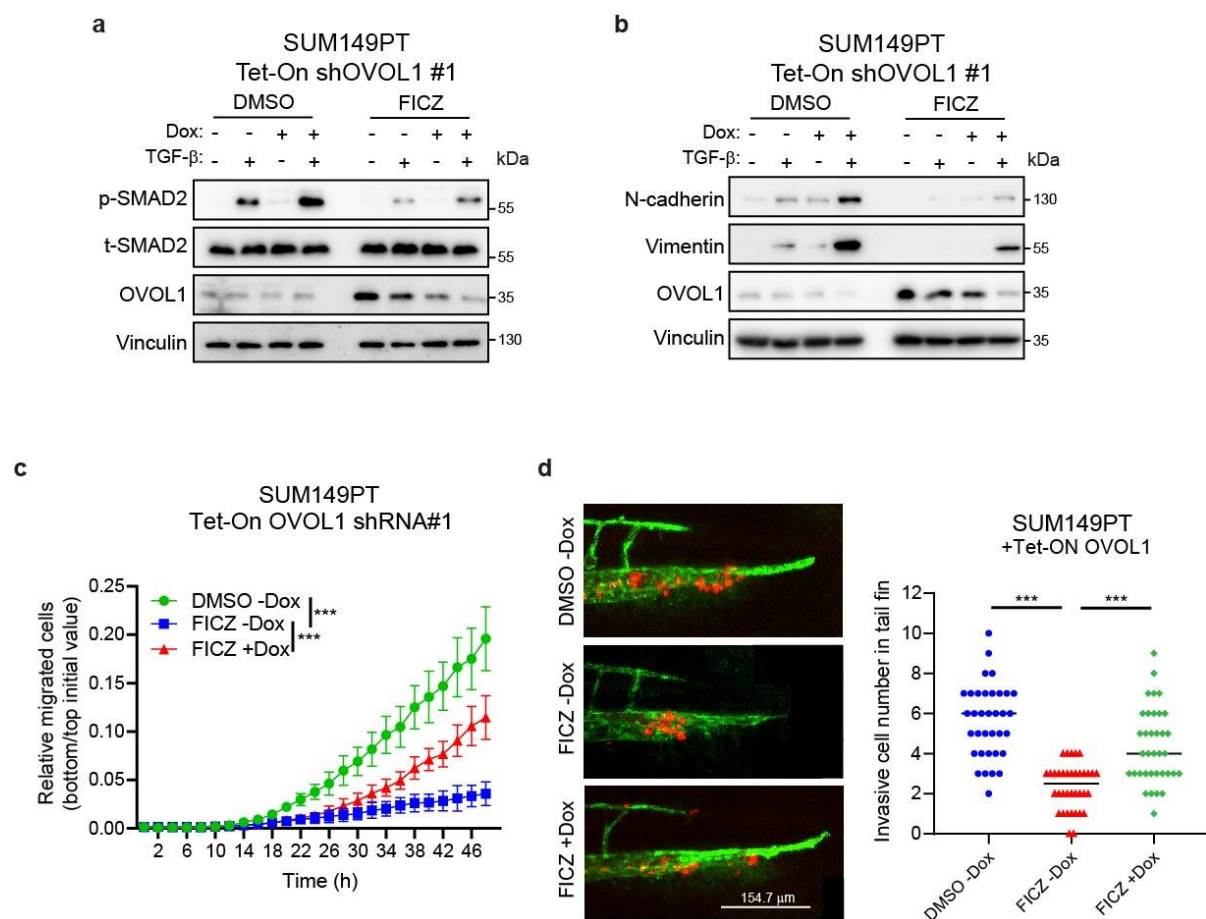
and OVOL1-FLAG in whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells transfected with MYC-SMAD7 and/or OVOL1-FLAG. **(d)** Western blotting detection of whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells transfected with FLAG-SMADs and OVOL1-MYC or OVOL2-MYC. **(e)** HA-SMAD7 expression quantified by Western blotting in HEK293T cells with inducible OVOL1 ectopic expression (left panel). Cells were kept in the presence or absence of Doxycycline (Dox) for 1 day, followed by the transfection of HA-SMAD7. Quantification of the relative protein level of HA-SMAD7 is shown in the right panel. To control for equal loading Vinculin levels were analyzed. The results are expressed as mean \pm SD. $** 0.001 < p < 0.01$. **(f, g)** Western blotting analysis of whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells without or with ectopic expression of OVOL1. Cells were treated without or with Doxycycline (Dox) for 1 day and then transfected with HA-Ub and FLAG-SMAD7. Total ubiquitination **(f)**, K48-ubiquitination **(g; left panel)** or K63-ubiquitination **(g; right panel)** of SMAD7 was probed by indicated antibodies.



Supplementary Fig. 8 The effects of RNF12 on SMAD7 are diminished by OVOL1. **(a)** Western blot analysis of indicated proteins in whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells transfected with HA-SMAD7 and FLAG-RNF12, FLAG-ARKADIA, FLAG-USP26 or FLAG-OTUD1. **(b)** Western blotting analysis of whole cell lysates (Input) and immunoprecipitants derived from HEK293T cells without or with ectopic expression of OVOL1. Cells were treated without or with Doxycycline (Dox) for 1 day and then transfected with HA-Ub, MYC-RNF12 and FLAG-SMAD7. K48-ubiquitination of SMAD7 was probed. **(c)** Schematic model of the OVOL1-induced interference of the interaction between RNF12 and SMAD7.



Supplementary Fig. 9 The inhibition of OVOL1 on TGF- β pathway is dependent on the interaction with SMAD7. (a) Reporter assay for detecting the luciferase activity in HEK293T cells transfected with TGF- β -induced SMAD3/4-dependent CAGA-luc transcriptional reporter and SMAD7 or OVOL1. Cells were stimulated with vehicle control (-) or TGF- β (5 μ M). The results are expressed as mean \pm SD. * 0.01 < p < 0.05, ** 0.001 < p < 0.01. (b) RT-qPCR measurement of *SMAD7* expression in HEK293T cells transfected with non-targeting siRNA (siNT) or siRNA targeting SMAD7 (siSMAD7). The results are expressed as mean \pm SD. ** 0.001 < p < 0.01. (c) Real-time scratch assay of MDA-MB-231 cells with inducible OVOL1 expression by Doxycycline (Dox). Representative scratch wounds are shown at the end time point of the experiment. The regions of original scratches and the areas of migrating cells are colored in purple and yellow respectively. (d) Luciferase activity in HEK293T cells transfected with TGF- β -induced SMAD3/4-dependent CAGA-luc transcriptional luciferase reporter and empty vector control (Co.vec), FLAG tagged full-length OVOL1 (FL) or indicated OVOL1 truncation mutants (T1-T3). Cells were stimulated with vehicle control (-) or TGF- β (5 μ M). The results are expressed as mean \pm SD. ** 0.001 < p < 0.01, *** 0.0001 < p < 0.001. (e) RT-qPCR measurement of *PAI-1* and *CTGF* expression in MDA-MB-231. Cells transfected with empty vector control (Co.vec), FLAG tagged full-length OVOL1 (FL) or indicated OVOL1 truncation mutants (T1-T3) were serum starved overnight and stimulated without or with TGF- β (1 ng/ml) for 4 h. The results are expressed as mean \pm SD. ** 0.001 < p < 0.01, *** 0.0001 < p < 0.001. (f) Western blotting analysis of the phosphorylation of SMAD2 (p-SMAD2) and total SMAD2 (t-SMAD2) in MDA-MB-231 cells. Cells transfected with empty vector (Co.vec), FLAG tagged full-length OVOL1 (FL) or indicated OVOL1 truncation mutants (T1-T3) were serum starved overnight and stimulated with vehicle control (-) or TGF- β (1 ng/ml) for 2 h. Vinculin levels were analyzed to control for equal loading.



Supplementary Fig. 10 OVOL1 inhibits TGF- β /SMAD2 signaling and EMT, cell migration and extravasation of SUM149PT cells. (a) Western blotting measurement of the phosphorylation of SMAD2 (p-SMAD2) and OVOL1 expression in SUM149PT cells with inducible OVOL1 knockdown by shRNA #1. Cells were stimulated without or with Doxycycline (Dox) for 2 days, followed by FICZ (5 μ M) treatment in serum starvation overnight before adding vehicle control or TGF- β (1 ng/ml) for another 2 h. The vehicle control DMSO was included for FICZ. Vinculin levels were analyzed to control for equal loading. (b) Western blotting analysis of mesenchymal markers expression in SUM149PT cells with inducible OVOL1 knockdown by shRNA #1. Cells were stimulated without or with Doxycycline (Dox) for 2 days, followed by FICZ (5 μ M) treatment overnight before adding TGF- β (1 ng/ml) for 2 days. Vehicle control DMSO was included for FICZ. (c) InCuCyte real-time chemotaxis assay for evaluating the migration of SUM149PT cells with inducible OVOL1 knockdown by shRNA #1. Cells were pre-treated without or with Doxycycline (Dox) for 2 days before seeding into the inserts, followed by FICZ treatment. Vehicle control DMSO was included for FICZ. The results are expressed as mean \pm SD. *** 0.0001 < p < 0.001. (d) *In vivo* zebrafish extravasation assay of SUM149PT cells without or with knockdown by shRNA #1. MQ water, Dox (to enable induction of the expression of OVOL1 shRNA) or FICZ (1 μ M) was added to the egg water from the first day post-injection. Representative images with zoom-in of the tail fin area are shown in the left panel. Analysis of the extravasated cell numbers in indicated groups is shown in the right panel. The results are expressed as mean \pm SD. *** 0.0001 < p < 0.001.

Supplementary tables are online at <https://www.nature.com/articles/s41392-022-00944-w>.



Appendix

English Summary
Nederlandse Samenvatting
List of Publications
Curriculum Vitae
Acknowledgements



English Summary

In cancer cells, malfunction of transforming growth factor (TGF)- β signaling can promote migration and metastasis, in part through the induction of epithelial-mesenchymal transition (EMT). Although strategies targeting TGF- β signaling are being explored in clinical trials, the on-target side effects caused by long-term systemic TGF- β signaling inhibition limit the clinical approval of TGF- β targeted therapies in cancer patients. Therefore, unraveling the regulatory mechanisms of TGF- β signaling in cancer (and normal) cells may offer new opportunities to treat cancer patients.

Long non-coding RNAs (lncRNAs) are a class of transcripts without coding potential but some were found to have a pivotal role in regulating signal transduction pathways through various mechanisms. In **Chapter 2**, we performed transcriptomic profiling to screen for TGF- β -induced lncRNAs in breast cancer cells. Follow-up loss-of-function studies identified lncRNA Induced by TGF- β and Antagonizes TGF- β Signaling 1 (*LITATS1*) as a protector of epithelial cells to suppress TGF- β -induced EMT and invasive abilities of cancer cells. Mechanistically, *LITATS1* serves as a scaffold to enforce the interaction between TGF- β type I receptor (T β RI) and the SMAD specific E3 ubiquitin ligase 2 (SMURF2), leading to the increase of polyubiquitination and proteasomal degradation of T β RI. *LITATS1* can also sequester SMURF2 protein in the cytoplasm, thereby promoting its export from the nucleus. Analysis of patient samples showed that *LITATS1* expression correlates with a favorable survival outcome in breast and non-small cell lung cancer patients, highlighting the potential of *LITATS1* as a promising prognostic marker. Of note, reintroducing *LITATS1* into highly aggressive breast cancer cells mitigated their migration and extravasation, suggesting that *LITATS1* may be a therapeutic anti-cancer agent.

lncRNAs that activate TGF- β signaling in cancer cells may be explored as alternative therapeutic targets to selectively inhibit TGF- β signaling and TGF- β -induced EMT in cancer cells. In **Chapter 3**, we described how lncRNA Enforcing TGF- β Signaling 1 (*LETS1*) promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β RI-stabilizing mechanism. In this study, we demonstrated that TGF- β /SMAD-induced nuclear *LETS1* interacted with nuclear factor of activated T cells (NFAT5) to facilitate the transcription of orphan nuclear hormone receptor *NR4A1*. NR4A1 alleviates T β RI polyubiquitination and potentiates T β RI stability by facilitating inhibitory (I)-SMAD7 protein degradation, leading to an activation of TGF- β /SMAD signaling, TGF- β -induced EMT, and cancer cell migration and extravasation. Thus, we unraveled a novel mechanism by which TGF- β /SMAD signaling is fine-tuned at the receptor level through an unannotated lncRNA *LETS1*.

Ovo-like transcriptional repressor 1 (OVOL1) is a vital determinant of epithelial lineage and stimulator of mesenchymal-epithelial transition (MET). However, its interplay with TGF- β and bone morphogenetic protein (BMP) signaling is unclear. **Chapter 4** presents that BMP strongly induces the expression of OVOL1, which potentiates BMP signaling in turn. This positive feedback loop is achieved by OVOL1-mediated suppression of TGF- β /SMAD signaling. OVOL1 binds to inhibitory SMAD7 and displaces interaction with E3 ligases targeting SMAD7. OVOL1 thereby prevents the polyubiquitination and proteasomal degradation of SMAD7. As a consequence, T β RI is destabilized by OVOL1, resulting in the attenuation of TGF- β signaling and TGF- β -induced EMT, migration in breast cancer cells. In addition, we identified 6-formylindolo(3,2-b)carbazole (FICZ) as a small-molecule compound that can stimulate OVOL1 expression and thereby antagonize (at least in part) TGF- β -triggered EMT

and migration in breast cancer cells. Hence, we uncovered a mechanism by which OVOL1 interplays with TGF- β and BMP signaling and maintains breast cancer cell epithelial identity.

Taken together, we identified several novel modulators of TGF- β /SMAD signaling. We studied the role of these modulators in TGF- β -induced EMT and migration in breast and lung cancer cells, and elucidated the mechanisms by which they fine-tune TGF- β /SMAD signaling transduction. These studies contribute to a better understanding of the regulatory networks of TGF- β signaling and may offer new therapeutic potentials to target TGF- β signaling in patients with breast or lung cancer.

Nederlandse Samenvatting

In kankercellen kan het afwijkend functioneren van de transformerende groeifactor (TGF)- β -signalering migratie en metastase bevorderen, gedeeltelijk door de inductie van epitheliale-mesenchymale transitie (EMT). Hoewel strategieën gericht op manipulatie van TGF- β -signalering worden onderzocht voor betere behandeling van kankerpatiënten, beperken de ontarget bijwerkingen veroorzaakt door langdurige systemische TGF- β -signaleringsremming de klinische goedkeuring. Daarom kan het ontrafelen van de regulerende mechanismen van TGF- β -signalering in kankercellen (en normale cellen) nieuwe mogelijkheden bieden om kankerpatiënten te behandelen.

Lange niet-coderende RNA's (lncRNA's) zijn een klasse transcripten zonder coderingspotentieel, maar sommige blijken een cruciale rol te spelen bij het reguleren van signaaltransductieroutes via verschillende mechanismen. In **Hoofdstuk 2** hebben we RNA expressie profilering uitgevoerd om te screenen op TGF- β -geïnduceerde lncRNAs in borstkankercellen. Vervolgstudies naar functieverlies identificeerden lncRNA geïnduceerd door TGF- β en antagoniseert TGF- β Signaling 1 (*LITATSI*) als een beschermer van epitheelcellen om de door TGF- β geïnduceerde EMT te onderdrukken. Mechanistisch gezien dient *LITATSI* als een platform om de interactie tussen TGF- β type I-receptor (T β RI) en de SMAD-specifieke E3 ubiquitineligase 2 (SMURF2) te versterken, wat leidt tot de toename van polyubiquitinatie en proteasomale afbraak van T β RI. *LITATSI* kan ook de locatie van SMURF2-eiwit in het cytoplasma stimuleren. Analyse van kankerpatiëntenmonsters toonde aan dat *LITATSI*-expressie correleert met een gunstig overlevingsresultaat bij borst- en niet-kleincellige longkankerpatiënten. Dit laatste benadrukt het potentieel van *LITATSI* als een veelbelovende prognostische kankermarker. De herintroductie van *LITATSI* in zeer agressieve borstkankercellen remt de migratie en invasie, wat suggereert dat *LITATSI* een therapeutisch middel tegen kanker zou kunnen zijn.

lncRNA's die TGF- β -signalering in kankercellen activeren, kunnen worden onderzocht als alternatieve therapeutische doelen om selectief TGF- β -geïnduceerde EMT in kankercellen te remmen. In **Hoofdstuk 3** hebben we beschreven hoe lncRNA Enforcing TGF- β Signaling 1 (*LETSI*) TGF- β -geïnduceerde EMT en kankeremigratie bevordert door transcriptieel een T β RI-stabiliserend mechanisme te activeren. In deze studie hebben we aangetoond dat door TGF- β /SMAD geïnduceerde nucleaire *LETSI* een interactie aangaat met de nucleaire factor van geactiveerde T-cellen (NFAT5) om de transcriptie van de nucleaire hormoonreceptor *NR4A1* te stimuleren. *NR4A1* remt de polyubiquitinatie van T β RI en versterkt de stabiliteit van T β RI door remmende (I)-SMAD7-eiwitafbraak te faciliteren. Dit leidt tot een activering van TGF- β /SMAD-signalering, door TGF- β geïnduceerde EMT en migratie en invasie van kankercellen. Zo hebben we een nieuw mechanisme ontrafeld waarmee TGF- β /SMAD-signalering op receptorniveau wordt verfijnd via een niet-geannoteerd lncRNA *LETSI*.

Ovo-achtige transcriptionele repressor 1 (OVOL1) is een cruciale bepalende factor voor de epitheliale cel identiteit en stimulator van de mesenchymale-epitheliale transitie (MET). **Hoofdstuk 4** laat zien dat botmorfogenetische proteïne (BMP) de expressie van OVOL1 sterk induceert, wat op zijn beurt de BMP-signalering versterkt. Deze positieve feedback wordt bereikt door OVOL1-gemedieerde onderdrukking van TGF- β /SMAD-signalering. OVOL1 bindt zich aan het remmende SMAD7 en verdringt de interactie met E3-ligasen die zich op SMAD7 richten. OVOL1 voorkomt daardoor de polyubiquitinatie en proteasomale afbraak van SMAD7. Als gevolg hiervan wordt T β RI gdestabiliseerd door OVOL1, wat resulteert in de verzwakking van TGF- β -signalering en TGF- β -geïnduceerde EMT-migratie in

borstkankercellen. Daarnaast hebben we 6-formylindolo(3,2-b)carbazol (FICZ) geïdentificeerd die de expressie van OVOL1 kan stimuleren en daardoor (tenminste gedeeltelijk) TGF- β -getriggerde EMT en migratie bij borstkanker kan tegenwerken. We hebben met deze bevindingen een mechanisme blootgelegd hoe OVOL1 de epitheliale identiteit van borstkankercellen handhaaft door TGF- β - en BMP-signalering te beïnvloeden.

Alles bij elkaar hebben we verschillende nieuwe modulators van TGF- β /SMAD-signalering geïdentificeerd. We bestudeerden de rol van deze modulators in TGF- β -geïnduceerde EMT en migratie in borst- en longkankercellen, en onderzochten de mechanismen waarmee ze TGF- β /SMAD-signaleringstransductie verfijnen. Deze onderzoeken dragen bij aan een beter begrip van de regulerende netwerken van TGF- β -signalering en kunnen nieuwe therapeutische mogelijkheden bieden voor patiënten met borst- of longkanker door heel gericht op TGF- β -signalering in kankercellen in te grijpen.

List of Publications

1. **Fan C**, Wang Q, Kuipers TB, Cats D, Iyengar PV, Hagenaars SC, Mesker WE, Devilee P, Tollenaar RAEM, Mei H, ten Dijke P. LncRNA *LITATSI* suppresses TGF- β -induced EMT and cancer cell plasticity by potentiating T β RI degradation. *The EMBO Journal*. 2023 May 15;42(10):e112806.
2. **Fan C**, González-Prieto R, Kuipers TB, Vertegaal ACO, van Veelen PA, Mei H, ten Dijke P. The lncRNA *LETS1* promotes TGF- β -induced EMT and cancer cell migration by transcriptionally activating a T β R1-stabilizing mechanism. *Science Signaling*, 2023 Jun 20;16(790):eadf1947.
3. **Fan C**, Wang Q, van der Zon G, Ren J, Agaser C, Slieker RC, Iyengar PV, Mei H, ten Dijke P. OVOL1 inhibits breast cancer cell invasion by enhancing the degradation of TGF- β type I receptor. *Signal Transduction and Targeted Therapy*. 2022 Apr 29;7(1):126.
4. Liu B, Wu T, Lin B, Liu X, Liu Y, Song G#, **Fan C**#, Ouyang G#. Periostin-TGF- β feedforward loop contributes to tumour-stroma crosstalk in liver metastatic outgrowth of colorectal cancer. *British Journal of Cancer*. 2024 Feb ;130(3):358-368. (#corresponding authors)
5. **Fan C***, Zhang J*, Hua W, ten Dijke P. Biphasic role of TGF- β in cancer progression: from tumor suppressor to tumor promoter. *Encyclopedia of Cancer (3rd Edition)*, Elsevier. 2017. (*co-first authors)
6. **Fan C***, Lin Y*, Mao Y*, Huang Z*, Liu AY, Ma H, Yu D, Maitikabili A, Xiao H, Zhang C, Liu F, Luo Q, Ouyang G. MicroRNA-543 suppresses colorectal cancer growth and metastasis by targeting KRAS, MTA1 and HMGA2. *Oncotarget*. 2016 Apr 19;7(16):21825-39. (*co-first authors)
7. **Fan C***, Lin B*, Huang Z*, Cui D*, Zhu M, Ma Z, Zhang Y, Liu F, Liu Y. MicroRNA-873 inhibits colorectal cancer metastasis by targeting ELK1 and STRN4. *Oncotarget*. 2018 Jan 2;10(41):4192-4204. (*co-first authors)
8. Lin Y*, Liu AY*, **Fan C***, Zheng H, Li Y, Zhang C, Wu S, Yu D, Huang Z, Liu F, Luo Q, Yang CJ, Ouyang G. MicroRNA-33b Inhibits Breast Cancer Metastasis by Targeting HMGA2, SALL4 and Twist1. *Scientific Reports*. 2015 Apr 28;5:9995. (*co-first authors)
9. Zhang Y, Xu G, Liu G, Ye Y, Zhang C, **Fan C**, Wang H, Cai H, Xiao R, Huang Z, Luo Q. miR-411-5p inhibits proliferation and metastasis of breast cancer cell via targeting GRB2. *Biochemical and Biophysical Research Communications*. 2016 Aug 5;476(4):607-613.
10. Sinha A, Mehta P, **Fan C**, Zhang J, Marvin DL, van Dinther M, Ritsma L, Boukany PE, ten Dijke P. Visualizing Dynamic Changes During TGF- β -Induced Epithelial to Mesenchymal Transition. *Methods in Molecular Biology*. 2022;2488:47-65.

Curriculum Vitae

Chuannan Fan was born on January 30th 1992 in Gaomi, Shandong province, China. In 2010 he obtained his bachelor degree. From September 2010, he started his bachelor study in biotechnology at the College of Life Sciences in Northwest A&F University, China. He obtained his bachelor degree in June 2014. From September 2014, Chuannan started his master study in cell biology at the College of Life Sciences in Xiamen University, China. During his master internship, under the supervision of Prof. Dr. Gaoliang Ouyang, he investigated functions and mechanisms of microRNAs and extracellular matrix proteins in cancer metastasis. He obtained his master degree in June 2017. From September 2017, Chuannan started his PhD study funded by a CSC scholarship at the Department of Cell and Chemical Biology department, Leiden University Medical Center, the Netherlands. During his PhD, he studied TGF- β signaling in cancer cells under the supervision of Prof. Dr. Peter ten Dijke. He focused on obtaining new insights of the underlying mechanisms that modulate TGF- β -induced epithelial to mesenchymal transition in cancer cells.

Acknowledgements

After six years of PhD study, I'm excited to show my research in this book. I really enjoy the PhD journey with the support from the colleagues, friends and family.

In particular, I would like to express my sincere gratitude to my supervisor Prof. Dr. Peter ten Dijke, who guided me to the gorgeous TGF- β world and showed me the beauty of signaling transduction and cell biology. With your profound knowledge and great patience, I got so much support and valuable feedback in the lab and in your office after fruitful discussions. I also want to express my appreciation to my co-advisors Dr. Manuel A. F. V. Gonçalves, Dr. David Baker and Dr. Prasanna Iyengar who gave me valuable comments and professional support in the experiments.

It's my great pleasure to be a member of our TGF- β family. Many thanks Maarten, Midory and Gerard for their enormous technical assistance in the lab. Thanks our current team members Agustin, Gunja, Chao, Weixin, Pranav, Kun and Jiying for their daily help and discussions during group meetings. Thanks our previous colleagues Jing, Jin Ma, Dieuwke, Wan, Jiang, Sijia, Laila, Maureen, Catalina, Marten, Abhishek, Sudha, Yongsheng, Yifan, Lizhe and Yuanzhuo for discussions and suggestions. Special thanks to Xinxin and Haijiang! Thanks my Chinese friends Jin Gan, Yufeng and Xueying for sharing their experience and experimental materials. Thanks Hans for valuable discussions and jokes. Thanks Babak, Arnaud, Jin Liu, Zhen, Xiaoling, Minkang and Shaozhuo.

Many thanks my collaborators Roderick, Davy, Tom, Leon (Hailiang), Román, Peter A van Veelen, Sophie, Wilma, Prof. Peter Devilee, and Prof. Rob A E M Tollenaar. I could not have completed the chapters in this thesis without your help!

Thanks to everyone in the CCB department. Thanks our secretaries Julia and Pauline. Thanks our trouble-shooter Willem. Thanks Annelies, Martijn, Gellila, Ben and Steve.

Last, special gratitude to my wife Qian, who offers me endless support during daily life and in the lab and fill my heart with love. Your encouragement, delicious food and our little boy Qi-an Fan (Dundun) make every day wonderful. I cannot complete this PhD adventure without you! Thanks my parents, who give me support and love in all the past years of my life!

I will remember all the memorable moments in this beautiful land and continue the journey!