A novel cancer-associated cassette exon in *TLN1* alters Talin 1 mechanosensitivity

Lina M. Gallego-Paez[1†], William J.S. Edwards[2†], Yanyu Guo[3], Manasa Chanduri[4], Thijs Koorman[5], Chieh-Yu Lee[1], Nina Grexa[1], Patrick W. B. Derksen[5], Jie Yan[3,6], Martin A. Schwartz[4,7], Jan Mauer[1,8]* and Benjamin T. Goult[2]*

[1]Biomed X Institute (GmbH), Heidelberg, Germany
[2]School of Biosciences, University of Kent, Canterbury CT2 7NJ, Kent, UK
[3]Mechanobiology Institute, National University of Singapore, Singapore 117411,
[4]Yale Cardiovascular Research Center and Departments of Internal Medicine (Cardiology)
[5]Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands
[6]Department of Physics, National University of Singapore, Singapore 117546, Singapore
[7]Departments of Cell Biology and Biomedical Engineering, Yale School of Medicine, New Haven CT 06511
[8]Current address: Department of Autoimmunity Transplantation and Inflammation, Novartis Institutes for BioMedical Research, Basel, Switzerland.

[†] These authors contributed equally to this work

* To whom correspondence should be addressed. Email: B.T.Goult@kent.ac.uk (Ben Goult) and jan.mauer@novartis.com (Jan Mauer)

Keywords: talin, integrin, splicing, splice variant, TLN1, CLSTN1, mechanotransduction, TGF-β

Abstract

Talin 1 is the core mechanosensitive adapter protein linking integrins to the cytoskeleton. The *TLN1* gene is comprised of 57 exons that encode the 2541 amino acid TLN1 protein. TLN1 was previously considered to be expressed as a single isoform. However, through differential pre-mRNA splicing analysis, we discovered a cancer-enriched, non-annotated 51-nucleotide exon in *TLN1* between exons 17 and 18, which we refer to as exon 17b. TLN1 is comprised of an N-terminal FERM domain, linked to 13 force-dependent switch domains, R1-R13. Inclusion of exon 17b results in an in-frame insertion of 17 amino acids immediately after Gln665 in the region between R1-R2 which lowers the force required to open the R1-R2 switches potentially altering downstream mechanotransduction. Biochemical analysis of this isoform revealed enhanced vinculin binding, and cells expressing this variant show altered adhesion dynamics. Finally, we show that the TGF-β/SMAD3 signaling pathway regulates this isoform switch. Future studies will need to consider the balance of these two TLN1 isoforms.

Introduction

Integrin-dependent adhesion of cells to the extracellular matrix (ECM) not only mediates cell attachment and physical integrity but is also critical in sensing the composition, organization and mechanical properties of the ECM. There are two talin genes in mammals, widely expressed Talin 1 (*TLN1*) and more specialized Talin 2 (*TLN2*) (Monkley et al., 2001; Senetar and McCann, 2005). Both Talins bind integrin beta subunit cytoplasmic domains, directly link to F-actin and thus, are essential for integrin-mediated adhesion (Calderwood et al., 2013; Klapholz and Brown, 2017). *TLN1* and *TLN2* have identical domain structure (Debrand et al., 2009; Gough and Goult, 2018), comprising an N-terminal head domain followed by an 80-residue unstructured linker and then a series of thirteen, 4- or 5- helical bundles designated rod domains 1-13 (R1-R13) (Goult et al., 2013). These helix bundles act as mechanochemical switches, that can open and close independently in response to mechanical force. Distinct protein partners bind specifically to helix bundles in the open or closed states such that domain opening and closing controls ligand binding (Goult et al., 2021, 2018). Force-dependent recruitment of signaling molecules thus enables talin to act as a mechanosensitive signaling hub between the ECM and the cell (Goult et al., 2018). As a result, talin plays a central role in cellular mechanosensing of ECM properties or applied strains. Talin-2 shows tissue-specific alternative splicing of the complex *TLN2* gene (Debrand et al., 2009). However, in studies to date, Talin 1 has been assumed to exist as a single constitutive isoform.

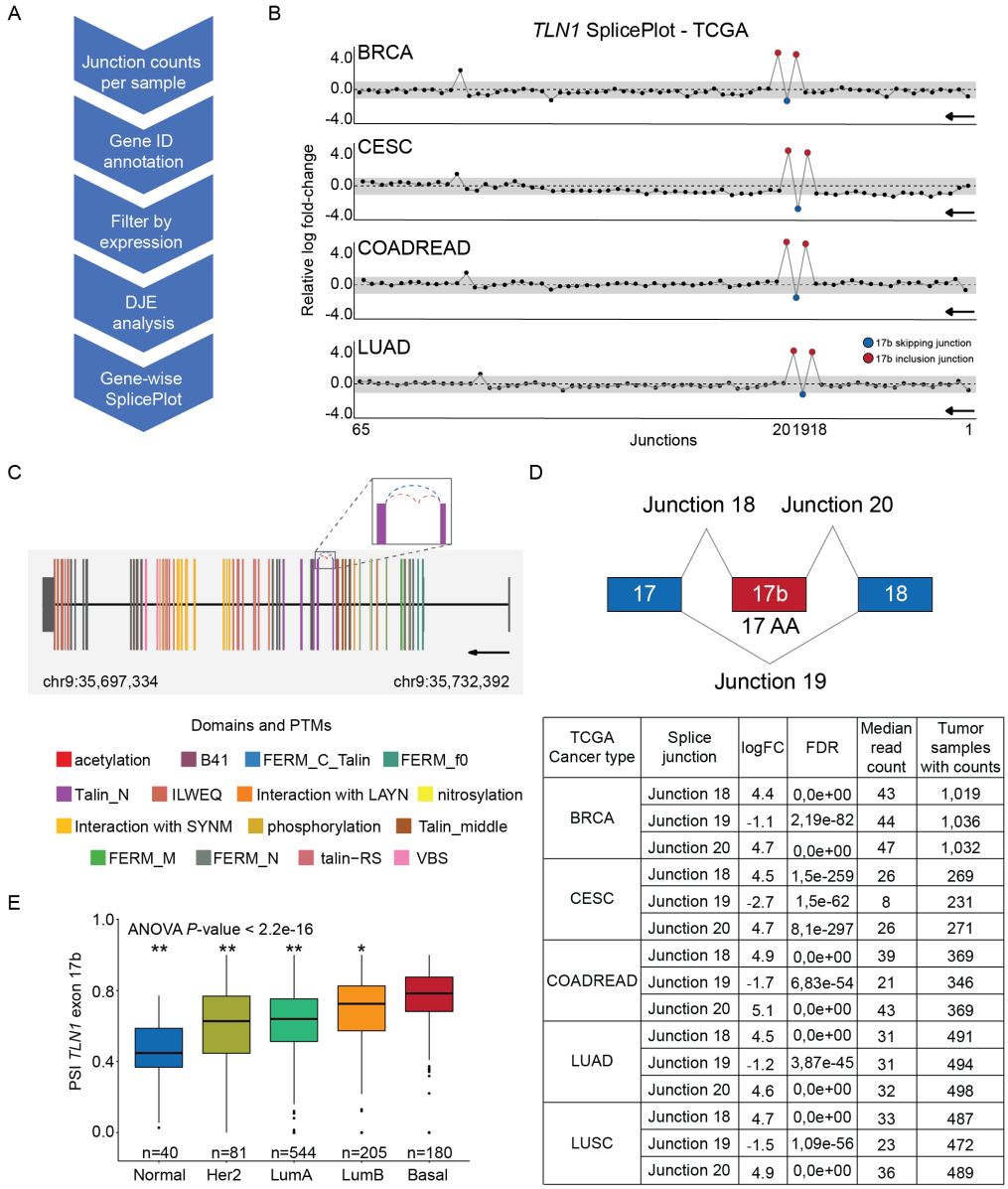
Bioinformatic pre-mRNA splicing analysis of pan-cancer RNA-Seq data sets from The Cancer Genome Atlas (TCGA) revealed a robustly transcribed undocumented mRNA sequence in *TLN1* (Gallego-Paez and Mauer, 2022). This region contains an additional non-annotated exon, located between exons 17 and 18, which we refer to as exon 17b. Sequence analysis revealed that exon 17b codes for an in-frame insertion of 17 amino acids into the first helix of the R2 helical bundle. Moreover, this novel *TLN1* exon is enriched in certain molecular cancer subtypes and is associated with altered drug responses and changes in gene dependencies in cancer cells. In this report, we characterized this novel splice variant of Talin and show that the *TLN1* gene is more complex than originally thought. Inclusion of the 17b cassette exon into *TLN1* mRNA provides a previously unrecognized mechanism for the cell to alter mechanotransduction in response to signals.

Results

Discovery of a novel exon in *TLN1*

We recently developed a splicing analysis pipeline, *DJExpress*, that uses junction expression information from RNA sequencing data to identify alternative splicing events in a

transcriptome-wide manner (Fig. 1A) (Gallego-Paez and Mauer, 2022). Notably, one of the features of DJExpress is the quantification of both annotated and non-annotated junctions. Thus, DJExpress supports the discovery of undocumented splice events.



TCGA Cancer type	Splice junction	logFC	FDR	Median read count	Tumor samples with counts
BRCA	Junction 18	4.4	0,0e+00	43	1,019
	Junction 19	-1.1	2,19e-82	44	1,036
	Junction 20	4.7	0,0e+00	47	1,032
CESC	Junction 18	4.5	1,5e-259	26	269
	Junction 19	-2.7	1,5e-62	8	231
	Junction 20	4.7	8,1e-297	26	271
COADREAD	Junction 18	4.9	0,0e+00	39	369
	Junction 19	-1.7	6,83e-54	21	346
	Junction 20	5.1	0,0e+00	43	369
LUAD	Junction 18	4.5	0,0e+00	31	491
	Junction 19	-1.2	3,87e-45	31	494
	Junction 20	4.6	0,0e+00	32	498
LUSC	Junction 18	4.7	0,0e+00	33	487
	Junction 19	-1.5	1,09e-56	23	472
	Junction 20	4.9	0,0e+00	36	489

Figure 1. **Differential splicing of a non-annotated *TLN1* cassette exon**. **(A)** The schematic shows the workflow of the *DJExpress* differential splicing analysis pipeline (DJE = differential junction expression). **(B)** Discovery of a novel cassette exon in TLN1 through differential splicing analysis of cancer patients and healthy controls. Gene-wise SlicePlots of *TLN1* mRNA generated with the *DJExpress* pipeline facilitated identification of two differentially expressed non-annotated junctions indicating the presence of an exon inclusion event (junctions 18-20) between exon 17 and 18 in several cancer types. Representative SlicePlots from breast invasive carcinoma (BRCA), cervical squamous cell carcinoma (CESC), colon and rectal adenocarcinoma (COADREAD) and lung adenocarcinoma (LUAD) cancer tissue from the Cancer Genome Atlas (TCGA) vs normal tissue from the Genotype-Tissue Expression (GTEx) database are shown. The black arrow indicates the direction of *TLN1* transcription on the reverse strand. Numbers on the x-axis indicate the first, last and differentially expressed junctions in the gene. Circles represent all junctions that were detected across the full-length TLN1 transcript. Significantly down- and upregulated junctions with |logFC| above cut-

off and FDR < 0.05 are shown in blue and red, respectively. Junctions with FDR > 0.05 for absolute or relative logFC (or both) are shown in black. Grey area indicates the log-foldchange cut-off (|logFC| > 1.0). **(C)** *TLN1* gene model plot with exon-to-protein domain annotation and differentially spliced junctions marked as dashed arcs connecting upstream and downstream exons. Coding and untranslated region (UTR) exons are illustrated as long and short exons respectively. Colors within exonic regions indicate the presence of protein domains and/or post translational modifications (PTMs) annotated within the Prot2HG protein domain database (Stanek et al., 2020). Differentially expressed junctions in *TLN1* are located within the TLN1 coding region. **(D)** Schematic showing the alternatively spliced region of *TLN1.* The constitutive exons 17 and 18 are shown in blue. The presence of junctions 18 and 20 indicate the inclusion of the 17 amino acid (AA) cassette exon 17b (red) whereas the presence of junction 19 indicates exon 17b exclusion. The table shows a summary of *TLN1* differentially expressed junctions (tumor vs normal tissue) in 5 example tumor types from TCGA (BRCA, CESC, COADREAD, LUAD, and lung squamous cell carcinoma (LUSC); LogFC = log fold-change). **(E)** Boxplots showing percent spliced-in (PSI) values of *TLN1* exon 17b across PAM50 subtypes of TCGA BRCA tumor samples (* *P*-value ≤ 0.01; ** *P*-value ≤ 0.0001 vs. basal subtype, one-way ANOVA).

Analysis of differential junction expression in TCGA data across all cancers identified such a non-annotated splice event in *TLN1* mRNA. Two novel junctions in *TLN1* mRNA, 18 and 20, were significantly increased in several TCGA cancer patient cohorts compared to healthy control tissue from the Genotype-Tissue Expression (GTEx) (Lonsdale et al., 2013) database (Fig. 1B). In parallel, we detected significantly reduced expression of the previously thought constitutive junction 19 of *TLN1* mRNA (Fig. 1B). These three junctions are located within the TLN1 coding region (Fig. 1C). Utilization of junctions 18 and 20 leads to the inclusion of an in-frame 51-nucleotide sequence flanked by exons 17 and 18, which we term exon 17b. Upon inclusion, 17 amino acids are inserted between TLN1 R1 and R2 (Fig. 1D).

Notably, exon 17b is enriched in breast, cervix, colon and lung cancer as indicated by reduced expression of the exon 17b-skipping junction 19 and increased expression of exon 17b inclusion junctions 18 and 20 (Fig. 1D). Moreover, in PAM50 breast cancer molecular subtypes, we found TLN1 exon 17b inclusion to be lowest in normal-like and highest in basal-like tumors (Fig. 1E), suggesting that TLN1 alternative splicing is part of the phenotypic differences between these subtypes.

Taken together, we have discovered a novel cassette exon 17b in TLN1 that shows significantly elevated expression across different cancers compared to healthy tissue, and whose level might be useful for stratification of certain molecular subtypes of breast cancer.

TLN1 exon 17b is differentially expressed in cancer cell lines

We next addressed whether *TLN1* exon 17b splicing can also be detected in cancer cell lines. Analysis of expression data from the Cancer Cell Line Encyclopedia (CCLE) (Barretina et al., 2012) revealed that exon 17b status varied within cancer subtypes including breast, lung and colon cancer (Fig. 2A and Fig. S1A). To independently validate splicing of exon 17b, we next performed RT-PCR on cDNA generated from two breast cancer cell lines with

high exon 17b inclusion (BT20, ZR751) and two breast cancer cell lines with low 17b inclusion (MDA231, BT549), according to *DJExpress* analysis (Fig. 2A). Indeed, exon 17b-specific RT-PCR of high exon 17b cell lines resulted in a 234 bp amplicon, indicating exon 17b inclusion (Fig. 2B). In contrast, cell lines with a predicted low expression of exon 17b produced a 183 bp amplicon, which reflects exon 17b skipping (Fig. 2B). Furthermore, we tested the inclusion of exon 17b in a panel of 2D and 3D grown cancer cell lines which we had available in the lab (Fig. S2) and here we clearly could detect both forms in 2D grown cells, but in 3D grown cells we exclusively saw the 17b version. The presence of *TLN1* exon 17b in cancer cell lines can thus be detected by conventional RT-PCR, reflected by a clear shift in amplicon size, without the need for RNA-Seq analysis.

Together these data demonstrate the sensitivity of the *DJExpress* pipeline for robustly identifying and classifying differential splicing events. Moreover, our finding that TLN1 exon 17b is present in some cancer cell lines but absent in others suggests that TLN1 alternative splicing might alter cancer cell behavior.

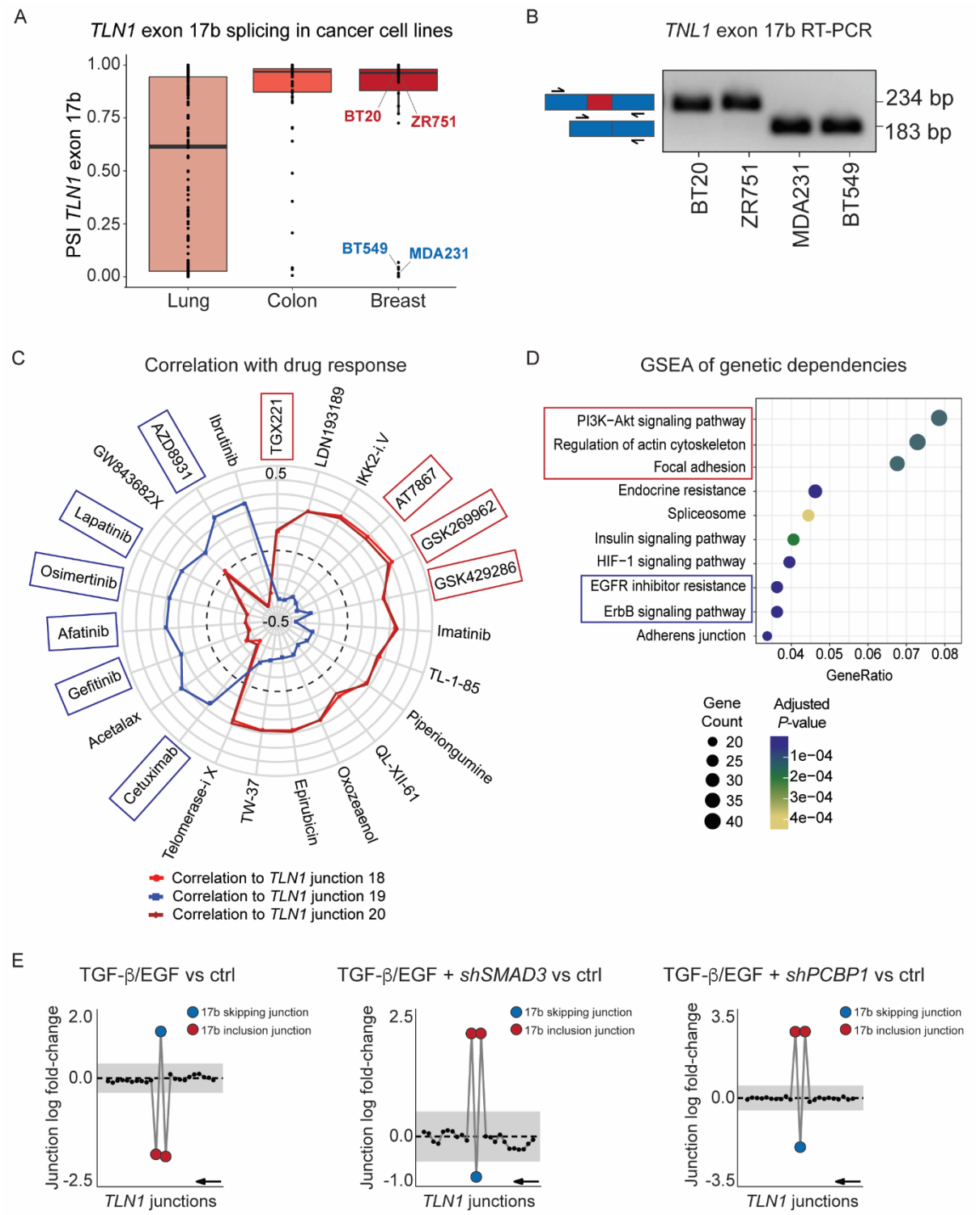


Figure 2. ***TLN1* exon 17b inclusion is associated with altered drug response and gene dependencies in cancer cell lines**. **(A)** Differential splicing of *TLN1* exon 17b is detected in cancer cell lines. Boxplots show distribution of percent spliced-in (PSI) values for *TLN1* exon 17b in lung, colon and breast cancer cell lines. The individually labelled BT20, ZR751, MDA231 and BT549 breast cancer cell lines were used for RT-PCR validation in (B). **(B)** RT-PCR validation of *TLN1* exon 17b expression in four representative breast cancer cell lines. RT-PCR was performed with primers flanking exon 17b (primer positions indicated by black arrows). Exon 17b spans 51 base pairs (bp) and exon 17b inclusion results in an amplicon size increase from 183 bp to 234 bp. BT20 and ZR751 cell lines show exon 17b inclusion whereas MDA231 and BT549 cell lines show exon 17b skipping. **(C)** Expression of *TLN1*

inclusion junctions (18, 20) and skipping junction (19) was correlated to cell survival after drug treatment using DepMap drug sensitivity data across all cancer cell lines. The top-ranked correlation coefficients (FDR < 0.05 and |rho| > 0.2) were used to construct the SpliceRadar plot. *TLN1* exon 17b inclusion (red and dark red lines) and exclusion (blue line) junction expression is plotted against their correlation coefficient with cell survival upon drug treatment. The data suggest that exon 17b inclusion is associated with increased sensitivity to EGFR inhibitors (blue boxes) and resistance to drugs targeting PI3K-Akt and cytoskeleton organization (red boxes). The black dashed line indicates a correlation coefficient *R* = 0 and an *R* range from -0.5 to 0.5 is shown. **(D)** KEGG gene set enrichment analysis (GSEA) of DepMap gene dependencies associated with *TLN1* exon 17b inclusion in cancer cell lines. The enrichment plot shows the top over-represented pathways, including cell adhesion, cytoskeleton organization (red) and EGFR/ErbB signaling pathways (blue). Dot size represents the number of genes enriched in each KEGG pathway and the color gradient indicates significance level of adjusted *P*-values. **(E)** Combined TGF-β/EGF treatment promotes TLN1 exon 17b skipping in a SMAD3 and PCBP1-dependent manner. Gene-wise splice plots of *TLN1* junction expression HeLa cells, which show baseline inclusion of exon 17b. Left panel: Combined TGF-β/EGF treatment leads to exon 17b skipping in HeLa cells. Middle panel: shRNA-mediated knockdown of TGF-β signal transducer SMAD3 blocks the TGF-β/EGF-induced skipping of exon 17b in HeLa cells. Right panel: shRNA-mediated knockdown of the RNA-binding protein PCBP1 blocks the TGF-β/EGF-induced skipping of exon 17b in HeLa cells. (The plots shown in this figure were generated by *DJExpress*-based re-analysis of RNA-Seq data from GSE72419; Grey area indicates the log-fold change cut-off (|logFC| > 0.5). Exon 17b inclusion junctions are shown in red, exon 17b skipping junction is shown in blue. Junctions with FDR > 0.05 for absolute or relative logFC (or both) are shown in black. Black arrow indicates the direction of *TLN1* transcription on the reverse strand.

TLN1 exon 17b is associated with altered drug responses and gene dependencies in cancer cell lines

We next asked whether exon 17b might functionally impact on cancer cell physiology. For this, we first tested whether exon 17b inclusion correlates with drug responses. Comparison of DepMap cancer cell line drug sensitivity data (Dempster et al., 2019; Corsello et al., 2019) with exon 17b inclusion and exclusion junction expression revealed that exon 17b inclusion correlates with increased sensitivity to EGFR inhibitors, represented by a negative correlation with drug area under the curve (AUC) (Fig. 2C). Conversely, cell lines that express TLN1 exon 17b were resistant to drugs targeting PI3K signaling and cytoskeleton regulation (Fig. 2C). Thus, exon 17b inclusion correlates with responses to a variety of drugs, suggesting that exon 17b status may support prediction of therapeutic success.

Therapeutic vulnerability can arise from an increased dependency on the expression of certain genes in cancer cells (Tsherniak et al., 2017). Since TLN1 exon 17b inclusion was associated with altered drug sensitivity, we reasoned that a similar association could also be detected for genetic dependencies. Thus, we next tested whether alternative splicing of *TLN1* is associated with altered gene dependencies in functional genomic screens. For this, we used pan-cancer cell line CRISPR screen data from the DepMap project and correlated exon 17b expression with gene effect. We then performed gene set enrichment analysis across all genes that significantly correlated with *TLN1* exon 17b inclusion. Here, we observed enrichment of pathways related to PI3K signaling, cytoskeleton organization, focal adhesion and EGFR/ErbB

signaling correlated with exon 17b expression (Fig. 2D). These data suggest that exon 17b status may differentially affect genetic dependencies in cancer cells and support the notion that exon 17b status alters cancer cell physiology. Notably, the enriched dependency pathways reflect the same pathways that are targeted by the drugs which we associated with 17b status, thus further supporting the notion that TLN1 alternative splicing links to changes in cancer cell behavior.

TLN1 exon 17b is regulated by TGF-β/SMAD3 signaling

Since nothing is known about the regulation of *TLN1* alternative splicing, we next wanted to identify regulators of exon 17b inclusion. For this, we re-analyzed published deep RNA-Seq data sets with DJExpress and assessed exon 17b status under different conditions. Re-analysis of a study that examined alternative splicing in HeLa cells, provided clear evidence of exon 17b status regulation (Tripathi et al., 2016). At baseline, HeLa cells show exon 17b inclusion, however stimulation with a combination of TGF-β and EGF lead to a clear splicing switch, which resulted in exon 17b skipping (Fig. 2E). Moreover, knockdown of the TGF-β downstream transcription factor SMAD3 blocked TGF-β/EGF-induced exon 17b skipping (Fig. 2E). Similar effects were observed after knockdown of the RNA binding protein PCBP1 (Fig. 2E). These results strongly suggest that the TGF-β-SMAD3-PCBP1 pathway regulates exon 17b splicing. Notably, TGF-β signaling is a well-described target in cancer therapy (Liu et al., 2021) and some of the beneficial effects of TGF-β inhibition drugs might be contributed to by changes in exon 17b status.

We next asked whether *TLN1* alternative splicing is correlated with other splicing events dysregulated in cancer. For this, we revisited our transcriptome-wide alternative splicing analysis in cancer vs. healthy tissue. Here, we found that splicing of exon 11 in Calsyntenin 1 (*CLSTN1*) negatively correlated with *TLN1* exon 17b splicing in lung and breast cancer patients and cell lines (Fig. S3A), suggesting that these two splice events might be "trans-mutually exclusive". This exclusivity was also reflected by RT-PCR analysis of breast cancer cell lines where cell lines with *TLN1* exon 17b inclusion exhibited *CLSTN1* exon 11 skipping and vice versa (Fig. S3B). Notably, *CLSTN1* exon 11 splicing exhibited the same dynamic, but inverse regulation upon TGF-β/EGF stimulation as *TLN1* exon 17b (Fig. S3C). These data indicate that *TLN1* and *CLSTN1* mRNA might be part of the same TGF-β/EGF-dependent alternative splicing program, which may be subject to future studies.

We next assessed *TLN1* exon 17b expression across all normal tissues by analysis of GTEx RNA-Seq data sets (Fig. S1B). Although expressed at low to non-detectable levels in many tissues, some tissues such as skin and pancreas show high expression of exon 17b and others

including kidney cortex, endocervix, testis, pituitary, liver and spleen show substantial levels as well. These data suggest that exon 17b is likely relevant to normal physiology.

Following our detection of this additional Talin 1 exon, we reported our finding to the ENSEMBL team who verified the presence of a missing cassette exon at chr9:35718356-35718406 (-), having extensive transcriptional support and being constrained as protein-coding in mammals. As a result, this exon was added to the new model ENST00000706939, which will appear in late 2022 in UniProt and ENSEMBL. Interestingly, there was evidence for exon 17b expression in some primates but exon 17b was to date not annotated in humans (Fig. S1C).

Exon 17b alters the primary sequence of Talin 1

We next addressed the possible impact of exon 17b on TLN1 protein function. TLN1 is comprised of 2541 amino acids, that make up the 18 talin domains (Fig. 3A). Exon 17b introduces a 17 amino acid, in-frame insertion immediately after residue Gln665 (Fig. 3B). Gln665 is in the first helix of the R2 bundle (Fig 3B-D). The crystal structure of talin R1R2 has been solved, which revealed that the two rod domains pack against each other in a side-to-side arrangement (Fig. 3C) (Papagrigoriou et al., 2004). To establish the effect of 17b on this region, we generated structural models of both R1R2-WT and the R1R2-17b proteins using the protein structure prediction tool AlphaFold (Senior et al., 2020). Both the R1R2-WT and R1R2-17b structural models showed good agreement with the crystal structure of the wildtype R1R2 (Papagrigoriou et al., 2004) validating that the models were accurate. The R1R2-17b structural model showed that the 17 amino acid insert extends the linker region between R1 and R2 (Fig. 3E, F) and is predominantly unstructured except for a small helical region in the linker (Fig. 3F), this helical region is from residues that were originally part of R2 before the insertion**.** The model also predicts structural differences within both the R1 and R2 domains. With insertion of the 17 aa sequence, the first helix of R2 is one turn shorter because the linker now incorporates the first three residues Pro662-Gln665, from the helix. Furthermore, in the wildtype R1R2 structure the linker between R1 and R2 is short, tightly linking the end of R2 with the start of R1 (Fig. 3D). The extended linker relieves this conformational constraint so that the last helix of R1 is extended by three residues (E657, S658, D659) (Fig. 3F). Interestingly, this extension comes from residues from in the R1R2 linker (blue) that, with the increased linker length, become part of the 5th helix of R1. It seems reasonable to assume that the helix could be extended in a similar fashion in WT-TLN1 if R2 was to unfold.

Full-length TLN1 can adopt open (Fig. 3A) and autoinhibited (Fig. 3G) states; overlaying the structural model of R1R2-17b on the cryo-EM structure of autoinhibited monomeric TLN1 (Dedden et al., 2019) shows that the 17b insertion is exposed on the surface of the closed conformation. In the context of the 250 kDa full-length molecule, these additional 17 residues

will be hard to detect by SDS-PAGE, which probably contributed to its late identification. In the context of R1R2 alone, the splice variant is readily detectible by SDS-PAGE due to the additional ~2 kDa (Fig. 3H). Together, these data suggest that the 17b insertion may perturb the mechanical behavior and function of the R1-R2 region of talin1.

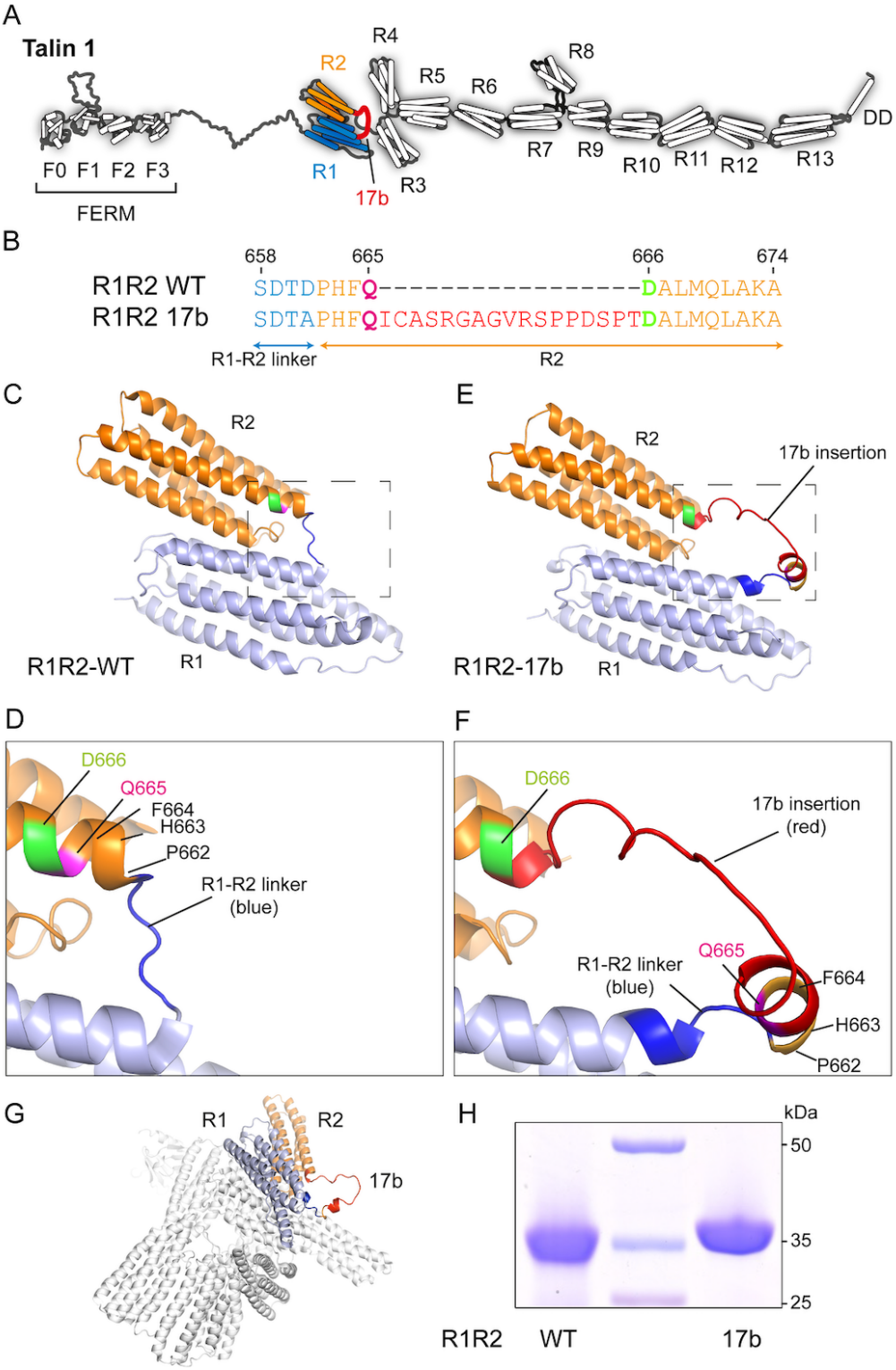


Figure 3. **Structural analysis of exon 17b at the protein level. (A)** Cartoon showing the location of the 17b insertion (red) in the TLN1 protein between R1 (blue) and R2 (orange). **(B)** Sequence alignment

of the region of R1R2-WT and R1R2-17b, showing the 17-residue insertion (red) in the R2 region (orange), the residues either side of the insertion, Q665 and D666 are highlighted in magenta and green respectively. **(C-D)** Crystal structure of R1R2-WT (Protein Data Bank accession no. 1SJ8). The same color scheme as in (B) is used with R1 (light blue), R1-R2 linker (blue) and R2 (orange). **(D)** Zoomed in region of the R1-R2 linker (blue) showing the location of the insertion site between Q665 (magenta) and D666 (green) in the first helix of R2. **(E-F)** AlphaFold model of R1R2-17b. **(E)** The 17b insertion is shown in red. **(F)** Zoomed in view of the same region as in (D). Residues on either side of the 17b insertion, Q665 (magenta) and D666 (green) are highlighted. The last helix of R1 is extended to contain part of the blue linker region, but the first helix of R2 is shortened. **(G)** Cryo-EM structure of full-length TLN1 (Protein Data Bank accession no. 6R9T) overlaid with the structural model of R1R2-17b, indicates that the insertion extends out of the autoinhibited monomeric state. **(H)** SDS-PAGE of the R1R2-WT and R1R2-17b recombinant proteins.

Exon 17b alters the biophysical and biochemical properties of R1R2

To investigate the biophysical and biochemical impacts of the 17-residue insertion on the R1R2 structure, we generated expression constructs, termed R1R2-WT and R1R2-17b. Both proteins expressed well and were amenable to characterization (Fig. 3H).

Far-UV CD spectra of R1R2-WT and R1R2-17b confirmed that both constructs were predominantly alpha-helical as predicted by structural analysis (Fig.3 and data not shown). However, CD thermostability analysis, where the CD is measured at a fixed wavelength (222 nm) over a range of temperatures revealed a significant difference in the stabilities of the two proteins (Fig. 4B). R1R2-WT showed a single transition confirming a cooperative unfolding event where both R1 and R2 helical bundles unfold simultaneously at a melting temperature (T_m) of 77°C. In contrast, R1R2-17b showed similar cooperative unfolding but a markedly reduced T_m of 64°C. This data reveals that the two switch domains R1 and R2 are both folded but destabilized by the 17b insertion.

Impact of 17b on the interactions of R1R2 with RIAM and vinculin

We next wanted to assess the effect of the 17b inclusion on the interactions of R1R2 with ligands. The Rap1-effector, RIAM binds to the folded R2 domain (Goult et al., 2013) while vinculin can bind to three vinculin binding sites (VBS); R1 contains a single VBS on helix 4 and R2 contains two VBS, in helices 1 and 4 of the 4-helix bundle, thus, fully unfolded R1R2 can bind up to 3 vinculins (Gingras et al., 2005) that are only available after mechanical force induces domain unfolding.

The stability of the bundle in which a VBS is embedded is critical as the domain needs to unfold to enable vinculin to bind, suggesting that the altered stability we observed via CD (Fig. 4A) might alter the VBS availability. To test how 17b effected vinculin binding we used analytical gel filtration where the talin proteins are preincubated with vinculin D1 domain (VD1) prior to loading on the size exclusion column. In line with earlier studies, only modest VD1 binding was seen with R1R2-WT after incubation at 37°C (Fig 4B) (Papagrigoriou et al., 2004).

However, substantially more complexation with R1R2-17b was observed, with both 1:2 and 1:3 talin:vinculin complex peaks detected. These data show that introduction of 17b reduces the stability of the R1R2 region and enhances vinculin binding.

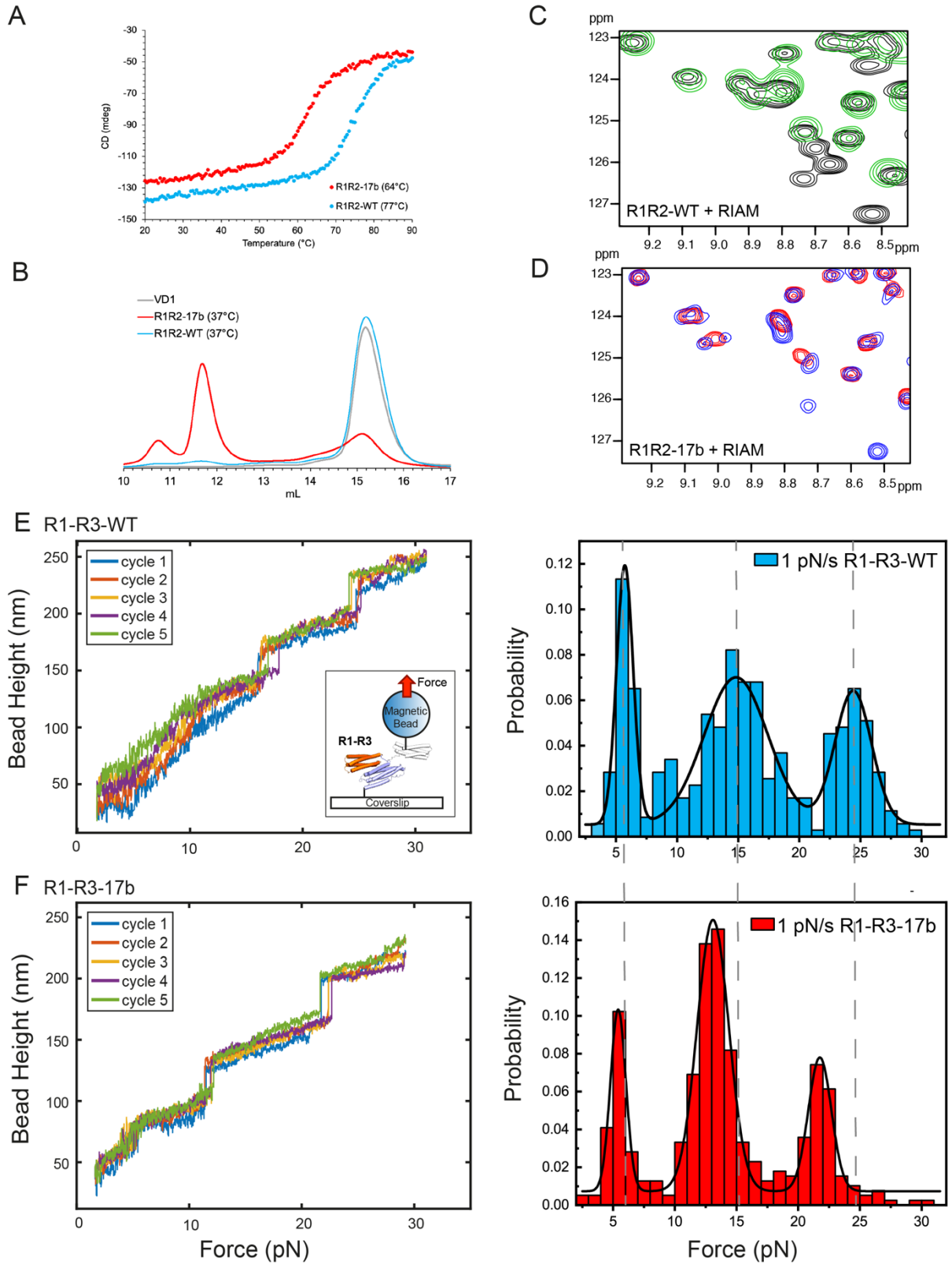


Figure 4**. Exon 17b alters the biophysical, biochemical and mechanical properties of R1R2. (A)** Circular dichroism (CD) analysis using proteins at 16 μM. Denaturation profiles for R1R2-WT (blue) and R1R2-17b (red) were measured by monitoring the change in CD at 222 nm with increasing temperature. Wild-type R1R2 has a melting temperature (T_m) of 77°C whereas R1R2-17b unfolds at 64˚C. **(B-D)** Characterization of R1R2-WT and R1R2-17b and interactions with RIAM and vinculin. **(B)** Vinculin domain 1 (VD1) binding analyzed by size exclusion chromatography (SEC) using a Superdex-200 gel filtration column. VD1 was incubated at a 3:1 ratio with R1R2-WT (blue) or R1R2-17b (red) at 37°C. Minimal binding was observed with the R1R2-WT. However, extensive complexation was observed for R1R2-17b and both 1:2 and 1:3 complex peaks were observed. **(C-D)** RIAM binding to R1R2 analyzed by NMR. $^{1}H,^{15}N$ TROSY spectra of 100 μM R1R2-WT (black) and R1R2-17b (blue) alone and upon addition of 300 μM RIAM peptide (green in (C), red in (D)). Extensive chemical shift changes were observed in both confirming RIAM interacts with both versions. **(E-F) Single-molecule stretching experiments show that 17b inclusion alters the mechanical stability of R1 and R2.** The experimental set up is shown in the inset of Panel E. A single molecule of R1-R3-WT (E) and R1-R3-17b (F) is tethered between a paramagnetic microbead and a coverslip. *Left panels*, Representative time traces of the height of the microbead during repeated force-increase scans at a loading rate of 1 pN s^{-1}. Each stepwise bead height increase indicates unfolding of a domain. Five independent force-increase scans are shown by the different color traces. *Right panels,* Normalized histograms of the unfolding forces of R1-R3-WT (N=353, seven independent tethers) and those of R1-R3-17b (N=391, nine independent tethers). The dotted lines indicate the unfolding thresholds for each of the domains in the wildtype R1-R3.

We next used NMR to assess RIAM binding by collecting NMR spectra of R1R2 alone, or in the presence of a 3-fold excess of RIAM peptide. Comparison of the $^{1}H,^{15}N$ TROSY spectra of the R1R2-WT and R1R2-17b showed both versions had similar peak dispersion, linewidths and peak positions, confirming that the core folding of the two domains was the same. Upon addition of a 3-fold excess of RIAM extensive chemical shift changes, predominantly to the peaks from the R2 domain were observed (Fig. 4C) indicative of binding as shown previously (Goult et al., 2013). Similar spectral changes were observed with R1R2-17b (Fig. 4D) indicating that the R1R2-17b variant binds similarly.

The effect of exon 17b on the mechanical response of talin

Talin mediates force transmission from ECM-bound integrin to cytoskeleton, and mechanotransduction via force-dependent interactions of talin-binding factors. To test whether the 17b inclusion alters the mechanical stability of R1 and R2, we performed single molecule experiments using our in-house developed magnetic-tweezer setup (Chen et al., 2011). In this assay, the protein is tethered between the microscope coverslip and a paramagnetic microbead and the magnetic tweezers are used to apply forces which increase with time at a fixed loading rate of 1 pN/s to the protein. The forces at which mechanical unfolding of the domains occurred (stepwise bead height increases in Fig.4E,F) are recorded over multiple cycles and with multiple tethers. For the single molecule experiments we used the R1-R3 triple domain construct as the mechanical response of this protein has been extensively characterized (Yao et al., 2014, 2016).

Both R1-R3-WT and R1-R3-17b show three unfolding events, with both having unfolding of a domain around the same force peaked at ~5 pN (Fig. 4E,F), which we assign to unfolding of

the R3 domain based on its highly characteristic signature reported in our previous studies (Yao et al., 2014, 2016). The unfolding of R2 and R1 domains occurs at higher force, which in the R1-R3-WT are the unfolding events at ~15 pN and 25 pN (Fig.4D). Strikingly, the introduction of 17b results in a marked reduction of the force required to unfold R2 and R1, which peaked at ~13 pN and ~21 pN (Fig.4F). Such a decrease by a few pN is significant since each unfolding step is coupled to large unfolding step sizes (>20 nm). Overall, these results indicate that the 17b inclusion results in a reduced mechanical stability of R1 and R2.

Cells expressing Talin1-17b have altered adhesions

To determine whether exon 17b affects TLN1 function in cells, *Tln1*$^{-/-}$ mouse embryonic fibroblasts were transfected with full-length TLN1 WT (Kumar et al., 2016) or full-length TLN1 exon 17b splice variant. Cells were then plated on fibronectin-coated coverslips and focal adhesion morphology was assessed by microscopy and quantified. Cells expressing TLN1 17b contained many more, smaller nascent adhesions as compared to WT Talin 1-expressing cells (Fig. 5A-C). However, the ratio of Vinculin to Talin within adhesions was equivalent (Fig. 5D). These data suggest that the TLN1 exon 17b variant alters the number and size of focal adhesions, with the net outcome that there are less mature focal adhesions.

Given the reduced stability of the R1R2-17b regions (Fig. 4), we also hypothesized that cells might have an altered morphological response to surfaces of different stiffness, which controls force on Talin (Kumar et al., 2016) . Stiffness sensing often examines spread cell area, which increases with increasing stiffness. However, when WT and 17b Talin 1-expressing cells were plated on fibronectin-coated polyacrylamide gels of variable stiffness, no difference in cell spread area was detected (Fig. 5E).

Together, these cellular experiments support a role for TLN1 exon 17b in altering focal adhesion formation but surprisingly this altered adhesion did not lead to changes in cellular substrate sensing, at least in the conditions tested. It is possible that the effects of this variant are less evident in baseline conditions i.e., without an additional stimulus.

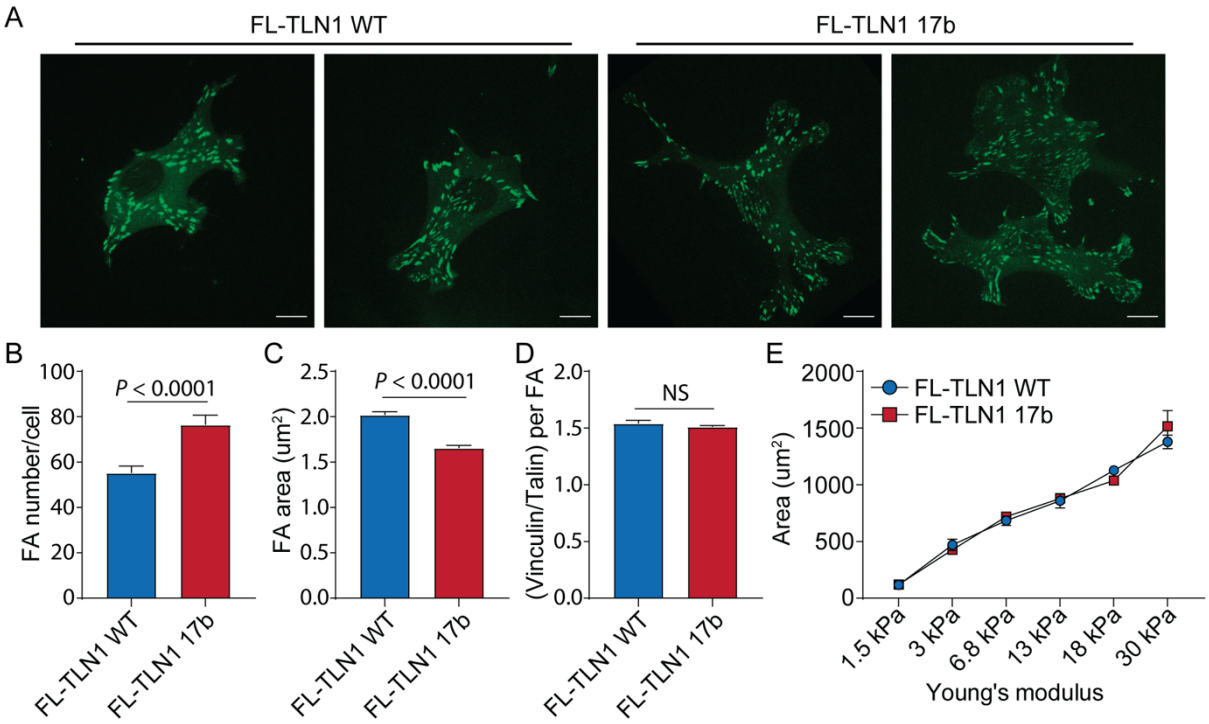


Figure 5. **The TLN1 exon 17b isoform alters focal adhesions in cells. (A)** Representative images of full-length Talin 1 WT and exon 17b splice variant expressed in Tln1$^{-/-}$ MEFs. Scale bar, 10 µm. Quantification of area **(B)**, number **(C)** and Vinculin to Talin 1 ratio **(D)** of WT and 17b Talin 1 splice variant containing focal adhesions. Data are means ± S.E.M, N=4735 and 4509 (B), N=60 and 58 cells (C) and N=620 and 970 (D) for WT and 17b Talin 1 containing focal adhesions respectively. NS=Not significant. **(E)** Cell area of Tln1$^{-/-}$ MEFs, expressing either WT or 17b splice variant form of Talin 1, plated for 6 h on polyacrylamide gels of varying stiffness. Data are mean ± S.E.M. N=3.

Conclusions

In summary, we report the discovery of a novel cassette exon 17b in the *TLN1* gene, which alters the Talin 1 protein coding sequence. This cassette exon can already be detected in healthy tissue but is highly enriched in several subtypes of cancer. Moreover, splicing of *TLN1* exon 17b can be controlled by TGF-β signaling pathways.

Until now, TLN1 has been thought to have a single isoform, so this discovery of an alternatively spliced isoform has significant implications for our understanding of integrin signaling. TLN1 forms a mechanosensitive signaling hub on the inside of the integrin adhesion complexes (Goult et al., 2021) where the 13 force-dependent binary switches in talin open and close in response to mechanical signals.

Whilst the full functional significance of Talin 1 alternative splicing is still to be determined, we show that the inclusion of this extra 17 amino acids markedly alters the stability of the R1 and R2 switches in the talin rod such that less force is required to open them. Modification of the mechanical stability of talin rod domain switches has been shown to alter the global mechanotransduction of the cell (Elosegui-Artola et al., 2014; Haining et al., 2018). The

consequence of this alteration in the R1 and R2 switches is that the binary patterns of information written into the shape of each talin molecule (Goult, 2021) will be different dependent on the status of the 17b exon, as different signaling outputs will be generated in response to the same amount of mechanical signals. Our recent analysis on the scale of talin domains unfolding (Barnett and Goult, 2022) revealed that helical bundle unfolding introduces a large extension into the length of the talin molecule, and R1R2 unfolding would introduce ~115 nm increase in length, relocating domains R3-R13 and all attached ligands further away from the membrane altering the spatial organization of molecules.

Exon 17b inclusion results in many more, smaller adhesions, but fewer larger adhesions. This phenotype is reminiscent to what we reported previously looking at nascent adhesion maturation (Han et al., 2021), where mutating the R8 domain of talin to reduce force-independent vinculin association resulted in many more nascent adhesions forming, but fewer maturing. It is tempting to speculate that the decreased stability of R1R2 and the concurrent increased vinculin binding to this region, results in altered adhesion dynamics by altering how force and cytoskeletal connections load onto the talin molecule maintaining adhesions in a more nascent state. It is also possible that the 17 amino acids, which extend out of both the open (Fig. 3A) and closed (Fig.3G) states of TLN1 present a novel binding motif for some, as of yet, unidentified ligand. Or it might be that the insertion, simply by altering the length and stability of the R1R2 domains, impacts on the mechanical response of talin and this is sufficient to exert global changes in cell signaling.

There now exist vast amounts of publicly available RNA-seq datasets from researchers looking at many different systems and different conditions, and this provides a wealth of information on this altered splicing event of *TLN1*. Our *DJExpress* pipeline enabled us to establish when and where the exon is included and that its inclusion is regulated by TGF-β/Smad3. TGF-β signaling is involved in many processes related to cancer malignancy and its use as a therapeutic target is actively investigated in the clinics (Hanahan and Weinberg, 2011; Derynck et al., 2021). The enrichment of *TLN1* exon 17b in several cancer subtypes together with the role of TGF-β signaling in its regulation thus could have far-reaching implications in the field of cancer biology.

Interestingly, we provide evidence that the *TLN1* splice variant is part of a concerted splicing program as we see inverse reciprocity between *TLN1* exon 17b inclusion and *CLSTN1* exon 11 exclusion. Further work would be required to fully appreciate the interplay between these previously unconnected proteins, but we speculate that they might form part of a change in the cellular program linked to the epithelial-mesenchymal transition (EMT).

Our finding that TLN1 exon 17b adhesions are smaller, but greater in number could indicate that upon certain stimuli, such as TGF-β, cells expressing the 17b variant could react more rapidly as smaller adhesions are inherently more dynamic in nature. Further exploration of these dynamics as well as identification of additional stimuli that regulate TLN1 splicing should be a priority for future studies.

We showed previously, how post-translation modification of TLN1 via cyclin-dependent kinase 1 (CDK1) can alter the mechanical response of the R7R8 region of talin, thereby changing the switch patterns that form (Gough et al., 2021). Here we show that alterations in the primary sequence of TLN1 via alternate splicing can also impact on the switch patterns, providing additional support to the emerging idea that mechanical signaling through TLN1 can be augmented via numerous chemical modifications providing a way to integrate mechanical and biochemical signals into signaling outputs. It will be important in future work to evaluate the status of the TLN1 isoforms being expressed in that cell type as alterations in cell signaling can alter mechanotransduction via modifying talin at the protein level.

Materials and methods

Human tumor and cell line datasets

RNA-seq data alignment, splice junction read quantification and sample quality-based filtering using 9,842 TCGA (Tomczak et al., 2015) tumor tissue samples across 32 different tumor types, 3,235 normal post-mortem tissue samples from GTEx (Lonsdale et al., 2013) and 1,019 cancer cell lines from the DepMap project (release 21Q3, www.depmap.org) was previously done as described in (Kahles et al., 2018; Gallego-Paez and Mauer, 2022). Clinical data associated to TCGA tumor samples was obtained using *TCGAbiolinks* package (Colaprico et al., 2016). Drug treatment and CRISPR screens data from cancer cell lines was retrieved from DepMap project portal (release 21Q3, www.depmap.org). RNA-seq data from TGF-β and EGF-treated HeLa cells was downloaded from the GEO database (accession number GSE72419) and raw data was processed with *DJExpress* (Gallego-Paez and Mauer, 2022).

Differential Junction Expression (DJE) analysis

Differential junction expression analysis was carried out using *DJExpress* tool (Gallego-Paez and Mauer, 2022). Briefly, quantified splice junction counts are gene-annotated, filtered by expression threshold (10 minimum read count mean per junction) and transformed to $\log_2$-counts per million (logCPM). Observation-level weights for each junction were computed using *limma* method (Ritchie et al., 2015) and linear modeling with empirical Bayes moderated *t*-statistics were implemented using *DJEanalyze* function to measure the significance level of observed changes in junction expression and usage between GTEx normal tissue and TCGA tumor tissue samples, between Fibroblasts “healthy” control cells and CCLE cancer cell lines and between TGF-β/EGF-treated and untreated HeLa cells. Gene-wise splice graphs and gene model plots with exon-to-protein domain annotation are produced using *DJEplotSplice* function.

Associations between TLN1 junction expression and functional genomics data

Significant linkages between *TLN1* exon 17b-related junction expression and drug treatment response or gene dependency values in cancer cell lines was carried out using correlation matrix operations implemented as part of the *DJEvsTrait* function from *DJExpress* (test.type = "Correlation", cor.method= "bicor", pvalue=0.05). *SpliceRadar* function was used for radar chart representation of top-ranked significant correlation coefficients. The clusterProfiler package version 4.4.2 (Yu et al., 2012) was used to identify enriched KEGG pathways within correlated gene dependencies to TLN1 junction expression, ranked by absolute correlation coefficient (adjusted *P*-value <0.05 for significantly enriched pathways).

Reverse transcription PCR (RT-PCR)

Total RNA was isolated from cancer cell lines and 1 µg was converted into cDNA using ProtoScript II First Strand cDNA Synthesis Kit (NEB). RT-PCR was performed with primers spanning the region between TLN1 exon 17 and exon 18 (TLN1_17b_fwd CAAGCAGCTGGGAACGTG, TLN1_17b_rev CTGAAGTCCCGAGTCCTCTG) and the region between CLSTN1 exon 10 and exon 12 (CLSTN1_11_fwd GACTCTCTATGTGGATGGCACG, CLSTN1_11_rev CCTTGCAGGTATACAGACAGTCG). The RT-PCR reactions were carried out at 94°C for 30 seconds and 94°C for 30 seconds, 55°C for 60 seconds, 68°C for 30 seconds for 30 cycles. The resulting PCR products were separated on a 1% agarose gel at 100 V for 1 h.

Cell line and organoid culture

BT20, BT549, MCF7, MCF10a, MDA-MB231, MDA-MB453, SKBR3, SUM44PE, T47D and ZR751 cell lines were obtained from the American Type Culture Collection (ATCC), STR typed verified by PCR and cultured as described previously. Derivation, establishment and culturing for the PDO models P008, P008::ΔCDH1 and KCL320 is extensively described in the supplementary materials of (Rätze et al., 2022). Culturing of PDO models HUB-72T (breast) and HUB-93T (breast) was published previously and are distributed by HUB ORGANOIDS (Sachs et al., 2018; Driehuis et al., 2019). Generation of E-cadherin knock-out MCF7::*ΔCDH1* and MCF10a::*ΔCDH1* using lentiviral CRISPR-Cas9 CDH1 targeted editing has been described previously in (Hornsveld et al., 2016). All cell lines and organoid models were routinely tested for Mycoplasma infection.

Protein expression and purification

Human Talin 1 R1R2-WT (residues 482-786) and R1R2-17b (residues 482-786 + additional 17 residues) were produced as codon optimized synthetic genes in pET151 plasmids (GeneArt). The vinculin Vd1 construct was also in pET151 as described previously (Wang et al., 2021). The RIAM peptide (residues 4-30; sequence SEDIDQMFSTLLGEMDLLTQSLGVDTC) was synthesized by GLBiochem. For the magnetic tweezers experiments human Talin 1 R1-R3-WT (residues 482-911) and R1-R3-17b (residues 482-911 + additional 17 residues) were cloned into a pGEX vector with Halo and Avi-tags as described in (Yao et al., 2016).

All constructs were transformed into BL21(DE3)* E. coli cells and grown in lysogeny broth + 100 µg/ml ampicillin at 37°C. Once OD had reached 0.7-0.8 the cultures were induced with 0.1 mM IPTG and incubated overnight at 20°C. Following harvesting the expressed proteins for biochemical analysis were purified using Nickel-NTA affinity chromatography. In short, after harvesting the cell pellets were resuspended in 20 mM Tris-HCl pH 8, 500 mM NaCl and 50

mM imidazole and lysed using sonication. Sonicated cell lysate was loaded into a 5 ml HisTrap HP column (GE Healthcare) and purified by nickel affinity chromatography. The eluted protein was dialyzed overnight into 20 mM Tris-HCl pH 8, 50 mM NaCl, during dialysis Invitrogen AcTEV protease was added to remove His-tags. Proteins were further purified using a HiTrap Q HP cation exchange column (GE Healthcare). Further details are available in (Khan et al., 2021). The GST-tagged constructs for the single-molecule studies were purified using glutathione Sepharose (GE Healthcare), and eluted by TEV cleavage.

The recombinant protein expression constructs have been deposited in Addgene at http://www.addgene.org/ben_goult.

AlphaFold structural modelling

To produce the structural models of human R1R2-WT and R1R2-17b, the 3D protein structure prediction tool, AlphaFold (Senior et al., 2020) was used. For this the sequence R1R2 +/- 17b was submitted to the software and the structural models visualized in PyMOL and compared against the crystal structure of mouse R1R2 (Protein Data Bank accession no. 1sj8 (Papagrigoriou et al., 2004)).

Circular dichroism (CD)

Circular dichroism was performed using a JASCO J-715 spectropolarimeter with 0.5 mg/ml protein samples. The far-UV spectra were collected between wavelengths of 200-260 nm, at 6 scans at a speed of 50 nm/min, 1 nm step resolution and a band width of 1 nm. CD melting curve data was collected at a wavelength of 222 nm, between 20-90°C, 1°C step resolution and 1 nm band width.

Size exclusion chromatography (SEC)

Size exclusion chromatography analysis of R1R2-WT, R1R2-17b and vinculin domain 1 (VD1) was done at room temperature. The total volume of samples was 100 µL, with protein concentrations between 100-400 µM (experiment dependent). The samples were loaded and ran using a Superdex 200 increase 10/300 GL. Preincubation at room temperature or 37°C was done for 1 h immediately prior to loading.

Nuclear magnetic resonance (NMR)

NMR samples were prepared in 50 mM NaCl, 15 mM NaH_2PO_4, 6 mM Na_2HPO_4, 2 mM DTT, pH 6.5, 5% (v/v) D_2O. All experiments were run at 298K on a Bruker Avance III 600 MHz NMR spectrometer equipped with CryoProbe. Data was processed with Topspin, and analyzed using CCPN Analysis (Skinner et al., 2015).

Single-molecule manipulation experiments

In the magnetic tweezer experiments, forces are applied to a protein of interest via superparamagnetic microbead and a 576 bp DNA linker using a custom magnetic tweezers platform that can exert forces up to 100 pN with ~1 nm extension resolution for tethered bead at 200 Hz sampling rate (Chen et al., 2011). The C-terminal of the protein of interest (R1-R3-17b or R1-R3-WT) was tethered to a Halo-ligand-coated coverslip via the C-terminal Halo-tag. The biotinylated N-terminal Avi-tag is linked to the biotinylated end of a 576 bp DNA linker via a traptavidin (Chivers et al., 2010). The other end of the DNA linker, which is labelled with a thiol group, is covalently attached to a superparamagnetic microbead (Dynabeads M270-epoxy). This way, a molecular tether consisting of the protein of interest and a DNA linker is spanned between the coverslip glass surface and the superparamagnetic microbead. Further

details can be found in our previous publication (Yao et al., 2016). All unfolding experiments were carried out in PBS, 3% BSA, 1 mM dithiothreitol and 0.1% Tween-20.
For given magnets and bead, the force is solely dependent on the magnet-bead distance $F(d)$, which can be calibrated based on a method described in our previous publication, which has an ~10% uncertainty due to the heterogeneous bead sizes (Chen et al., 2011). The force loading rate control is achieved by decreasing the magnet-bead distance $d(t)$ in a manner such that the force increases at a constant rate r (Zhao et al., 2017).

Mammalian expression plasmids, cell transfection and immunofluorescence

Full-length WT Talin 1 plasmid containing internal GFP and RFP (Kumar et al., 2016) was used in this study. Nucleotides corresponding to the 17-residue region of the 17b splice variant was cloned into WT Talin 1 by Gibson assembly. Briefly, the WT Talin 1 plasmid was digested with NotI and XhoI. Primers containing overhangs of nucleotides corresponding to the 17-residue region of the splice variant were used to amplify R2. Two fragments containing nucleotides between the NotI site to R1 domain and beyond R2 domain to XhoI site were also amplified by PCR. All fragments were incubated with Gibson assembly mix (New England Biolabs) as per manufacturer's instructions.

Tln1$^{-/-}$ cells (Priddle et al., 1998) were transiently transfected with full-length WT Talin 1 and full-length Talin 1-17b variant, using Lipofectamine 2000 (ThermoFisher Scientific) for 48 h. Cells were trypsinized and plated on either fibronectin-coated (10 µg/mL) glass-bottom dishes or polyacrylamide gels of varying stiffness (Aratyn-Schaus et al., 2010). At 6 h, cells were fixed in 4% paraformaldehyde and counterstained with Alexa 647 phalloidin (ThermoFisher Scientific). Mouse monoclonal antibody for vinculin (Sigma, V9131) was used to visualize vinculin. Cells were imaged using a 63X objective in a confocal microscope. ImageJ was used to assess cell area, focal adhesion area and focal adhesion number.

Acknowledgements

We thank Sridevi Jaksani for technical support with organoid culturing. B.T.G. was funded by Biotechnology and Biological Sciences Research Council (Grant BB/S007245/1) and Cancer Research UK Program Grant (DRCRPG-May21\100002). J.M. and L.M.G-P. were funded by Merck KGaA, Darmstadt, Germany (CrossRef Funder ID: 10.13039/100009945). M.S. was funded by USPHS grant (R01 GM047214-25). J.Y. was funded by The Ministry of Education under the Research Centres of Excellence programme and Singapore Ministry of Education Academic Research Funds Tier 2 (MOE-T2EP50220-0015). T.K. and P.W.B.D. received financial support from Breast Cancer Now (2018novPCC1297) which is supported by Pfizer, the European Union's FET Proactive program under the grant agreement No. 731957 (MECHANO-CONTROL), and COST action LOBSTERPOT (CA19138), supported by COST (European Cooperation in Science and Technology).

Author Contributions

L.M.G-P. performed the bioinformatic analyses with support from C-Y.L.. T.K., C-Y.L. and N.G. performed the RT-PCR analyses. W.J.S.E. performed the biochemical and biophysical

analysis and the computational modelling. Y.G. and J.Y. performed the single-molecule analysis. M.C. and M.S. performed the cell biology and stiffness sensing experiments. T.K. and P.W.B. provided conceptual input. J.M. and B.T.G. conceived the study, supervised the project and wrote the paper with input from all authors. All authors have read the manuscript and were given the opportunity to provide input.

Conflict of interest

At the time this study was conceived, L.M.G-P. and J.M, were employees of BioMed X Institute (GmbH), Heidelberg, Germany, and their research was funded by the healthcare business of Merck KGaA, Darmstadt, Germany (CrossRef Funder ID: 10.13039/100009945). The healthcare business of Merck KGaA, Darmstadt, Germany had no part in the study design and data collection, analysis or interpretation of the results but provided feedback regarding the general research strategy. J.M. is currently an employee at Novartis Pharma AG, Basel, Switzerland. Novartis Pharma AG had no part in this study.

References

Aratyn-Schaus, Y., P.W. Oakes, J. Stricker, S.P. Winter, and M.L. Gardel. 2010. Preparation of compliant matrices for quantifying cellular contraction. *J. Vis. Exp.* 46. doi:10.3791/2173.

Barnett, S.F.H., and B.T. Goult. 2022. The MeshCODE to scale – Visualising synaptic binary information. *bioRxiv*. 2022.06.16.496395. doi:10.1101/2022.06.16.496395.

Barretina, J., G. Caponigro, N. Stransky, K. Venkatesan, A.A. Margolin, S. Kim, C.J. Wilson, J. Lehár, G. V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M.F. Berger, J.E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jané-Valbuena, F.A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I.H. Engels, J. Cheng, G.K. Yu, J. Yu, P. Aspesi, M. De Silva, K. Jagtap, M.D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R.C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J.P. Mesirov, S.B. Gabriel, G. Getz, K. Ardlie, V. Chan, V.E. Myer, B.L. Weber, J. Porter, M. Warmuth, P. Finan, J.L. Harris, M. Meyerson, T.R. Golub, M.P. Morrissey, W.R. Sellers, R. Schlegel, and L.A. Garraway. 2012. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 483:603–607. doi:10.1038/NATURE11003.

Calderwood, D. a, I.D. Campbell, and D.R. Critchley. 2013. Talins and kindlins: partners in integrin-mediated adhesion. *Nat. Rev. Mol. Cell Biol.* 14:503–17. doi:10.1038/nrm3624.

Chen, H., H. Fu, X. Zhu, P. Cong, F. Nakamura, and J. Yan. 2011. Improved High-Force Magnetic Tweezers for Stretching and Refolding of Proteins and Short DNA. *Biophys. J.* 100:517–523. doi:10.1016/j.bpj.2010.12.3700.

Chivers, C.E., E. Crozat, C. Chu, V.T. Moy, D.J. Sherratt, and M. Howarth. 2010. A streptavidin variant with slower biotin dissociation and increased mechanostability. *Nat. Methods*. 7:391–393. doi:10.1038/nmeth.1450.

Colaprico, A., T.C. Silva, C. Olsen, L. Garofano, C. Cava, D. Garolini, T.S. Sabedot, T.M. Malta, S.M. Pagnotta, I. Castiglioni, M. Ceccarelli, G. Bontempi, and H. Noushmehr. 2016. TCGAbiolinks: An

R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44:e71. doi:10.1093/nar/gkv1507.

Corsello, S.M., R.T. Nagari, R.D. Spangler, J. Rossen, M. Kocak, J.G. Bryan, R. Humeidi, D. Peck, X. Wu, A.A. Tang, V.M. Wang, S.A. Bender, E. Lemire, R. Narayan, P. Montgomery, U. Ben-David, Y. Chen, M.G. Rees, N.J. Lyons, J.M. McFarland, B.T. Wong, L. Wang, N. Dumont, P.J. O'Hearn, E. Stefan, J.G. Doench, H. Greulich, M. Meyerson, F. Vazquez, A. Subramanian, J.A. Roth, J.A. Bittker, J.S. Boehm, C.C. Mader, A. Tsherniak, and T.R. Golub. 2019. Non-oncology drugs are a source of previously unappreciated anti-cancer activity. *bioRxiv*. 730119. doi:10.1101/730119.

Debrand, E., Y. El Jai, L. Spence, N. Bate, U. Praekelt, C.A. Pritchard, S.J. Monkley, and D.R. Critchley. 2009. Talin 2 is a large and complex gene encoding multiple transcripts and protein isoforms. 276:1610–1628. doi:10.1111/j.1742-4658.2009.06893.x.

Dedden, D., S. Schumacher, C.F. Kelley, M. Zacharias, C. Biertümpfel, R. Fässler, and N. Mizuno. 2019. The Architecture of Talin1 Reveals an Autoinhibition Mechanism. *Cell*. 179:120-131.e13. doi:10.1016/j.cell.2019.08.034.

Dempster, J.M., J. Rossen, M. Kazachkova, J. Pan, G. Kugener, D.E. Root, and A. Tsherniak. 2019. Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines. *bioRxiv*. 720243. doi:10.1101/720243.

Derynck, R., S.J. Turley, and R.J. Akhurst. 2021. TGFβ biology in cancer progression and immunotherapy. *Nat. Rev. Clin. Oncol.* 18:9–34. doi:10.1038/s41571-020-0403-1.

Driehuis, E., S. Kolders, S. Spelier, K. Lõhmussaar, S.M. Willems, L.A. Devriese, R. de Bree, E.J. de Ruiter, J. Korving, H. Begthel, J.H. van Es, V. Geurts, G.W. He, R.H. van Jaarsveld, R. Oka, M.J. Muraro, J. Vivié, M.M.J.M. Zandvliet, A.P.A. Hendrickx, N. Iakobachvili, P. Sridevi, O. Kranenburg, R. van Boxtel, G.J.P.L. Kops, D.A. Tuveson, P.J. Peters, A. van Oudenaarden, and H. Clevers. 2019. Oral mucosal organoids as a potential platform for personalized cancer therapy. *Cancer Discov.* 9:852–871. doi:10.1158/2159-8290.CD-18-1522.

Elosegui-Artola, A., E. Bazellières, M.D. Allen, I. Andreu, R. Oria, R. Sunyer, J.J. Gomm, J.F. Marshall, J.L. Jones, X. Trepat, and P. Roca-Cusachs. 2014. Rigidity sensing and adaptation through regulation of integrin types. *Nat. Mater.* 13:631–637. doi:10.1038/nmat3960.

Gallego-Paez, L.M., and J. Mauer. 2022. DJExpress: An Integrated Application for Differential Splicing Analysis and Visualization. *Front. Bioinforma.* 2:19. doi:10.3389/fbinf.2022.786898.

Gingras, A.R., W.H. Ziegler, R. Frank, I.L. Barsukov, G.C.K.K. Roberts, D.R. Critchley, and J. Emsley. 2005. Mapping and consensus sequence identification for multiple vinculin binding sites within the talin rod. *J. Biol. Chem.* 280:37217–37224. doi:10.1074/jbc.M508060200.

Gough, R.E., and B.T. Goult. 2018. The tale of two talins - two isoforms to fine-tune integrin signalling. *FEBS Lett.* 592:2108–2125. doi:10.1002/1873-3468.13081.

Gough, R.E., M.C. Jones, T. Zacharchenko, S. Le, M. Yu, G. Jacquemet, S.P. Muench, J. Yan, J.D. Humphries, C. Jørgensen, M.J. Humphries, and B.T. Goult. 2021. Talin mechanosensitivity is modulated by a direct interaction with cyclin-dependent kinase-1. *J. Biol. Chem.* 297:100837. doi:10.1016/j.jbc.2021.100837.

Goult, B.T. 2021. The Mechanical Basis of Memory-the MeshCODE Theory. *Front. Mol. Neurosci.* 14:1–18. doi:10.3389/fnmol.2021.592951.

Goult, B.T., N.H. Brown, and M.A. Schwartz. 2021. Talin in mechanotransduction and mechanomemory at a glance. *J. Cell Sci.* 134. doi:10.1242/JCS.258749.

Goult, B.T., J. Yan, and M.A. Schwartz. 2018. Talin as a mechanosensitive signaling hub. *J. Cell Biol.* 217:3776–3784. doi:10.1083/jcb.201808061.

Goult, B.T., T. Zacharchenko, N. Bate, R. Tsang, F. Hey, A.R. Gingras, P.R. Elliott, G.C.K. Roberts, C. Ballestrem, D.R. Critchley, and I.L. Barsukov. 2013. RIAM and vinculin binding to talin are mutually exclusive and regulate adhesion assembly and turnover. *J. Biol. Chem.* 288:8238–8249. doi:10.1074/jbc.M112.438119.

Haining, A.W.M., R. Rahikainen, E. Cortes, D. Lachowski, A. Rice, M. von Essen, V.P. Hytönen, and A. del Río Hernández. 2018. Mechanotransduction in talin through the interaction of the R8 domain with DLC1. *PLOS Biol.* 16:e2005599. doi:10.1371/journal.pbio.2005599.

Han, S.J., E. V. Azarova, A.J. Whitewood, A. Bachir, E. Guttierrez, A. Groisman, A.R. Horwitz, B.T. Goult, K.M. Dean, and G. Danuser. 2021. Pre-complexation of talin and vinculin without tension is required for efficient nascent adhesion maturation. *Elife*. 10. doi:10.7554/eLife.66151.

Hanahan, D., and R.A. Weinberg. 2011. Hallmarks of cancer: The next generation. *Cell*. 144:646–674. doi:10.1016/j.cell.2011.02.013.

Hornsveld, M., M. Tenhagen, R.A. Van De Ven, A.M.M. Smits, M.H. Van Triest, M. Van Amersfoort, D.E.A. Kloet, T.B. Dansen, B.M. Burgering, and P.W.B. Derksen. 2016. Restraining FOXO3-dependent transcriptional BMF activation underpins tumour growth and metastasis of E-cadherin-negative breast cancer. *Cell Death Differ.* 23:1483–1492. doi:10.1038/cdd.2016.33.

Kahles, A., K. Van Lehmann, N.C. Toussaint, M. Hüser, S.G. Stark, T. Sachsenberg, O. Stegle, O. Kohlbacher, C. Sander, S.J. Caesar-Johnson, J.A. Demchok, I. Felau, M. Kasapi, M.L. Ferguson, C.M. Hutter, H.J. Sofia, R. Tarnuzzer, Z. Wang, L. Yang, J.C. Zenklusen, J. (Julia) Zhang, S. Chudamani, J. Liu, L. Lolla, R. Naresh, T. Pihl, Q. Sun, Y. Wan, Y. Wu, J. Cho, T. DeFreitas, S. Frazer, N. Gehlenborg, G. Getz, D.I. Heiman, J. Kim, M.S. Lawrence, P. Lin, S. Meier, M.S. Noble, G. Saksena, D. Voet, H. Zhang, B. Bernard, N. Chambwe, V. Dhankani, T. Knijnenburg, R. Kramer, K. Leinonen, Y. Liu, M. Miller, S. Reynolds, I. Shmulevich, V. Thorsson, W. Zhang, R. Akbani, B.M. Broom, A.M. Hegde, Z. Ju, R.S. Kanchi, A. Korkut, J. Li, H. Liang, S. Ling, W. Liu, Y. Lu, G.B. Mills, K.S. Ng, A. Rao, M. Ryan, J. Wang, J.N. Weinstein, J. Zhang, A. Abeshouse, J. Armenia, D. Chakravarty, W.K. Chatila, I. de Bruijn, J. Gao, B.E. Gross, Z.J. Heins, R. Kundra, K. La, M. Ladanyi, A. Luna, M.G. Nissan, A. Ochoa, S.M. Phillips, E. Reznik, F. Sanchez-Vega, C. Sander, N. Schultz, R. Sheridan, S.O. Sumer, Y. Sun, B.S. Taylor, J. Wang, H. Zhang, et al. 2018. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell*. 34:211-224.e6. doi:10.1016/j.ccell.2018.07.001.

Khan, R.B., L. Varela, A.R. Cowell, and B.T. Goult. 2021. Biochemical Characterization of the Integrin Interactome. *Methods Mol. Biol.* 2217:115–147. doi:10.1007/978-1-0716-0962-0_9.

Klapholz, B., and N.H. Brown. 2017. Talin – the master of integrin adhesions. *J. Cell Sci.* 130:2435–2446. doi:10.1242/jcs.190991.

Kumar, A., M. Ouyang, V. den D. K, E.J. McGhee, K. Tanaka, M.D. Anderson, A. Groisman, B.T. Goult, K.I. Anderson, and M.A. Schwartz. 2016. Talin tension sensor reveals novel features of focal adhesion force transmission and mechanosensitivity. *J. Cell Biol.* 213:371–383. doi:10.1083/jcb.201510012.

Liu, S., J. Ren, and P. ten Dijke. 2021. Targeting TGFβ signal transduction for cancer therapy. *Signal Transduct. Target. Ther.* 6. doi:10.1038/s41392-020-00436-9.

Lonsdale, J., J. Thomas, M. Salvatore, R. Phillips, E. Lo, S. Shad, R. Hasz, G. Walters, F. Garcia, N. Young, B. Foster, M. Moser, E. Karasik, B. Gillard, K. Ramsey, S. Sullivan, J. Bridge, H. Magazine, J. Syron, J. Fleming, L. Siminoff, H. Traino, M. Mosavel, L. Barker, S. Jewell, D. Rohrer, D. Maxim,

D. Filkins, P. Harbach, E. Cortadillo, B. Berghuis, L. Turner, E. Hudson, K. Feenstra, L. Sobin, J. Robb, P. Branton, G. Korzeniewski, C. Shive, D. Tabor, L. Qi, K. Groch, S. Nampally, S. Buia, A. Zimmerman, A. Smith, R. Burges, K. Robinson, K. Valentino, D. Bradbury, M. Cosentino, N. Diaz-Mayoral, M. Kennedy, T. Engel, P. Williams, K. Erickson, K. Ardlie, W. Winckler, G. Getz, D. DeLuca, Daniel MacArthur, M. Kellis, A. Thomson, T. Young, E. Gelfand, M. Donovan, Y. Meng, G. Grant, D. Mash, Y. Marcus, M. Basile, J. Liu, J. Zhu, Z. Tu, N.J. Cox, D.L. Nicolae, E.R. Gamazon, H.K. Im, A. Konkashbaev, J. Pritchard, M. Stevens, T. Flutre, X. Wen, E.T. Dermitzakis, T. Lappalainen, R. Guigo, J. Monlong, M. Sammeth, D. Koller, A. Battle, S. Mostafavi, M. McCarthy, M. Rivas, J. Maller, I. Rusyn, A. Nobel, F. Wright, A. Shabalin, et al. 2013. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* 45:580–585. doi:10.1038/ng.2653.

Monkley, S.J., C.A. Pritchard, and D.R. Critchley. 2001. Analysis of the mammalian talin2 gene TLN2. *Biochem. Biophys. Res. Commun.* 286:880–885. doi:10.1006/bbrc.2001.5497.

Papagrigoriou, E., A.R. Gingras, I.L. Barsukov, N. Bate, I.J. Fillingham, B. Patel, R. Frank, W.H. Ziegler, G.C. Roberts, D.R. Critchley, and J. Emsley. 2004. Activation of a vinculin-binding site in the talin rod involves rearrangement of a five-helix bundle. *EMBO J.* 23:2942–2951. doi:10.1038/sj.emboj.7600285.

Priddle, H., L. Hemmings, S. Monkley, A. Woods, B. Patel, D. Sutton, G.A. Dunn, D. Zicha, and D.R. Critchley. 1998. Disruption of the talin gene compromises focal adhesion assembly in undifferentiated but not differentiated embryonic stem cells. *J. Cell Biol.* 142:1121–1133. doi:10.1083/jcb.142.4.1121.

Rätze, M.A.K., T. Koorman, T. Sijnesael, B. Bassey-Archibong, R. van de Ven, L. Enserink, D. Visser, S. Jaksani, I. Viciano, E.R.M. Bakker, F. Richard, A. Tutt, L. O'Leary, A. Fitzpatrick, P. Roca-Cusachs, P.J. van Diest, C. Desmedt, J.M. Daniel, C.M. Isacke, and P.W.B. Derksen. 2022. Loss of E-cadherin leads to Id2-dependent inhibition of cell cycle progression in metastatic lobular breast cancer. *Oncogene*. 41:2932–2944. doi:10.1038/s41388-022-02314-w.

Ritchie, M.E., B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, and G.K. Smyth. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi:10.1093/nar/gkv007.

Sachs, N., J. de Ligt, O. Kopper, E. Gogola, G. Bounova, F. Weeber, A.V. Balgobind, K. Wind, A. Gracanin, H. Begthel, J. Korving, R. van Boxtel, A.A. Duarte, D. Lelieveld, A. van Hoeck, R.F. Ernst, F. Blokzijl, I.J. Nijman, M. Hoogstraat, M. van de Ven, D.A. Egan, V. Zinzalla, J. Moll, S.F. Boj, E.E. Voest, L. Wessels, P.J. van Diest, S. Rottenberg, R.G.J. Vries, E. Cuppen, and H. Clevers. 2018. A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. *Cell*. 172:373-386.e10. doi:10.1016/j.cell.2017.11.010.

Senetar, M.A., and R.O. McCann. 2005. Gene duplication and functional divergence during evolution of the cytoskeletal linker protein talin. *Gene*. 362:141–152. doi:10.1016/j.gene.2005.08.012.

Senior, A.W., R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A.W.R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D.T. Jones, D. Silver, K. Kavukcuoglu, and D. Hassabis. 2020. Improved protein structure prediction using potentials from deep learning. *Nature*. 577:706–710. doi:10.1038/s41586-019-1923-7.

Skinner, S.P., B.T. Goult, R.H. Fogh, W. Boucher, T.J. Stevens, E.D. Laue, and G.W. Vuister. 2015. Structure calculation, refinement and validation using CcpNmr Analysis. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 71:154–161. doi:10.1107/S1399004714026662.

Stanek, D., D.M. Bis-Brewer, C. Saghira, M.C. Danzi, P. Seeman, P. Lassuthova, and S. Zuchner. 2020. Prot2HG: A database of protein domains mapped to the human genome. *Database*. 2020:161. doi:10.1093/database/baz161.

Tomczak, K., P. Czerwińska, and M. Wiznerowicz. 2015. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Wspolczesna Onkol.* 1A:A68–A77. doi:10.5114/wo.2014.47136.

Tripathi, V., K.M. Sixt, S. Gao, X. Xu, J. Huang, R. Weigert, M. Zhou, and Y.E. Zhang. 2016. Direct Regulation of Alternative Splicing by SMAD3 through PCBP1 Is Essential to the Tumor-Promoting Role of TGF-β. *Mol. Cell*. 64:549–564. doi:10.1016/j.molcel.2016.09.013.

Tsherniak, A., F. Vazquez, P.G. Montgomery, B.A. Weir, G. Kryukov, G.S. Cowley, S. Gill, W.F. Harrington, S. Pantel, J.M. Krill-Burger, R.M. Meyers, L. Ali, A. Goodale, Y. Lee, G. Jiang, J. Hsiao, W.F.J. Gerath, S. Howell, E. Merkel, M. Ghandi, L.A. Garraway, D.E. Root, T.R. Golub, J.S. Boehm, and W.C. Hahn. 2017. Defining a Cancer Dependency Map. *Cell*. 170:564-576.e16. doi:10.1016/j.cell.2017.06.010.

Wang, Y., M. Yao, K.B. Baker, R.E. Gough, S. Le, B.T. Goult, and J. Yan. 2021. Force-dependent interactions between talin and full-length vinculin. *JACS*. (in press). doi:10.1101/2021.04.26.441533.

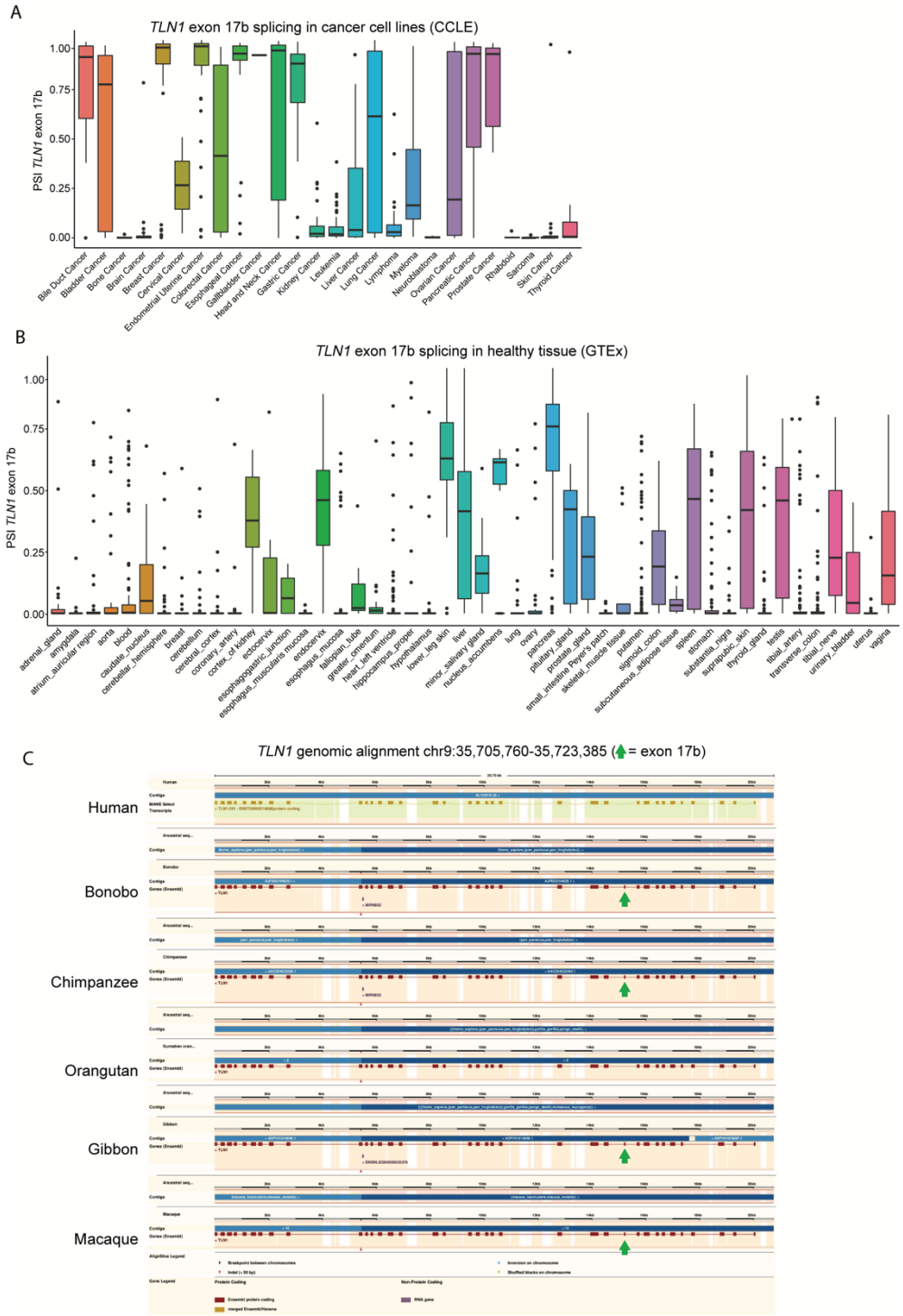
Yao, M., B.T. Goult, H. Chen, P. Cong, M.P. Sheetz, and J. Yan. 2014. Mechanical activation of vinculin binding to talin locks talin in an unfolded conformation. *Sci. Rep.* 4:4610. doi:10.1038/srep04610.

Yao, M., B.T. Goult, B. Klapholz, X. Hu, C.P. Toseland, Y. Guo, P. Cong, M.P. Sheetz, and J. Yan. 2016. The mechanical response of talin. *Nat. Commun.* 7:11966. doi:10.1038/ncomms11966.

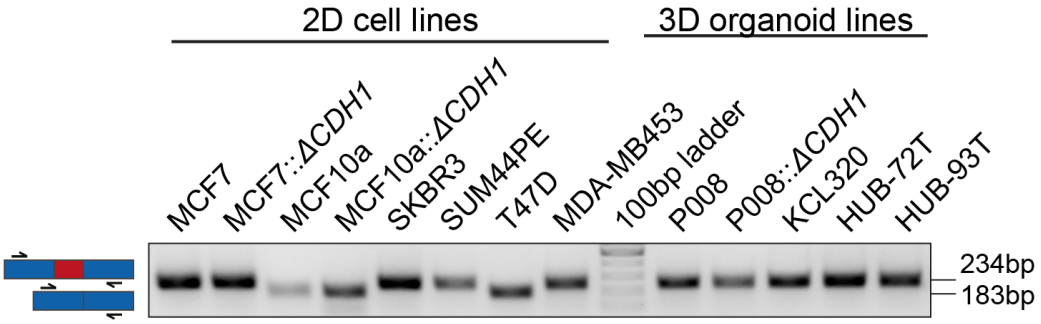
Yu, G., L.G. Wang, Y. Han, and Q.Y. He. 2012. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* 16:284–287. doi:10.1089/omi.2011.0118.

Zhao, X., X. Zeng, C. Lu, and J. Yan. 2017. Studying the mechanical responses of proteins using magnetic tweezers. *Nanotechnology*. 28. doi:10.1088/1361-6528/aa837e.

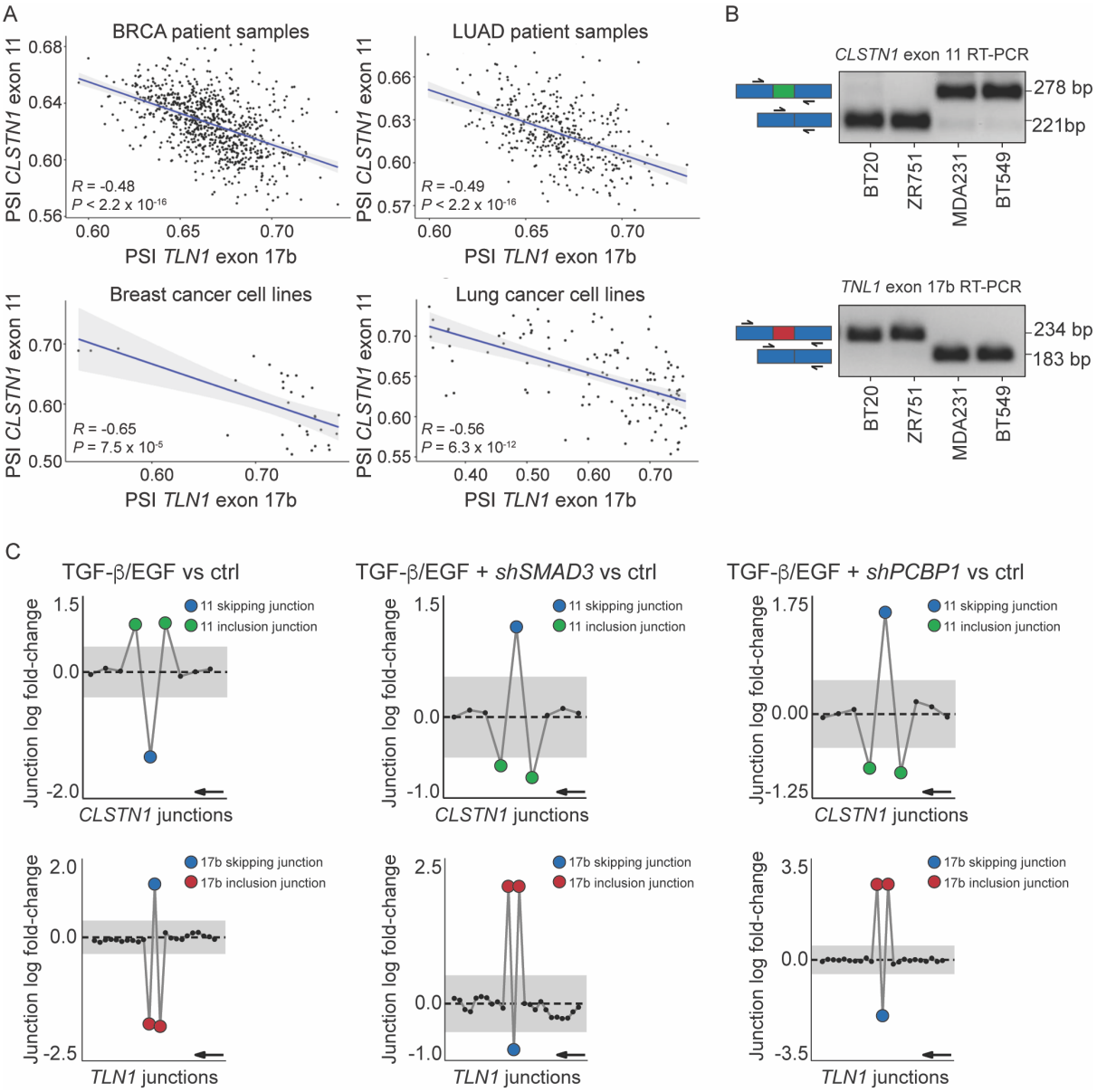
Supplementary Information.



Supplementary Figure 1. **The non-annotated exon 17b in *TLN1* mRNA is utilized across cancer cell lines and has variable expression levels in healthy tissues. (A)** Percent spliced-in (PSI) values of *TLN1* exon 17b in CCLE cancer subtypes. **(B)** PSI values of *TLN1* exon 17b across GTEx healthy tissues. Although absent in many healthy tissue types, exon 17b is highly enriched in some tissues including skin and pancreas. **(C)** Alignment of *TLN1* gene structure in humans and primates. The green arrows indicate the position of *TLN1* exon 17b, which is annotated in bonobo, chimpanzee, gibbon and macaque but not documented in humans and orangutan.



Supplementary Figure 2. ***TLN1* exon 17b inclusion/exclusion in cancer cell lines grown in 2D and 3D**. The *TLN1* exon 17b expression status in 2D breast cancer cell lines and 3D breast cancer organoid models was assessed using RT-PCR. RT-PCR was performed with primers flanking exon 17b (primer positions indicated by black arrows). Exon 17b spans 51 base pairs (bp) and exon 17b inclusion results in an amplicon size increase from 183 bp to 234 bp analyzed on an agarose gel. Note that the organoid models grown in 3D all express the 17b inclusion version of Talin.



Supplementary Figure 3. ***CLSTN1* exon 11 expression is inversely correlated with *TLN1* exon 17b expression in cancer patients and cell lines (A)** Correlation analysis of percent spliced-in (PSI) values of *TLN1* exon 17b and *CLSTN1* exon 11 in tumor patient tissue and cancer cell lines. *TLN1* exon 17b and *CLSTN1* exon 11 splicing is inversely correlated suggesting nearly trans-mutually exclusive splicing regulation of these two events (*R* = Pearson correlation coefficient, *P* = adjusted *P*-value). **(B)** RT-PCR validation of *CLSTN1* exon 11 expression in four representative breast cancer cell lines. RT-PCR was performed with primers flanking exon 11 (primer positions indicated by black arrows). Exon 11 skipping results in an amplicon size reduction from 278 bp to 221 base pairs. BT20 and ZR751 cell lines show exon 11 skipping whereas MDA231 and BT549 cell lines show exon 11 inclusion. The RT-PCR analysis of TLN1 exon 17b expression in cancer cell lines shown in Figure 2B was repurposed here to illustrate the inverse splicing pattern with *CLSTN1* exon 11. **(C)** Analogous analysis to Figure 2E, that reveals dynamic *CLSTN1* exon 11 splicing in response to combined TGF-β/EGF treatment. Gene-wise splice plots of *CLSTN1* junction expression in HeLa cells, which show baseline skipping of exon 11. The

analysis of the *TLN1* exon 17b expression in HeLa cells shown in Figure 2E was repurposed here to illustrate the inverse splicing pattern with *CLSTN1* exon 11 (The plots shown in this figure were generated by *DJExpress*-based re-analysis of RNA-Seq data from GSE72419; Grey area indicates the log-fold change cut-off (|logFC| > 0.5). Inclusion junctions are shown in red, skipping junctions are shown in blue. Junctions with FDR > 0.05 for absolute or relative logFC (or both) are shown in black. Black arrow indicates the direction of transcription on the reverse strand).