

Supplementary Information

**De novo identification of universal cell
mechanics regulators**

Urbanska, Ge, *et al.*

Supplementary Tables

Supplementary Table 1 | Combined *PC-corr* values calculated as means of the two analyzed sets specifying the values for network edges. Edges above a cut-off of 0.75 are displayed. The cut-off is indicated with a horizontal line at the bottom.

	node <i>i</i>	node <i>j</i>	edge, $PC-corr_{i,j}^{comb}$
1	<i>FHL2</i>	<i>THBS1</i>	0.863
2	<i>ANKRD1</i>	<i>IL11</i>	0.813
3	<i>MFAP5</i>	<i>THBS1</i>	0.803
4	<i>FHL2</i>	<i>IGFBP7</i>	0.788
5	<i>IGFBP7</i>	<i>THBS1</i>	0.788
6	<i>C1QTNF1</i>	<i>IGFBP6</i>	0.785
7	<i>FHL2</i>	<i>TAGLN</i>	0.782
8	<i>IGFBP7</i>	<i>TAGLN</i>	0.782
9	<i>TAGLN</i>	<i>THBS1</i>	0.782
10	<i>ATP8B1</i>	<i>FHL2</i>	0.780
11	<i>CNN2</i>	<i>FHL2</i>	0.774
12	<i>FHL2</i>	<i>MFAP5</i>	0.767
13	<i>LRRC15</i>	<i>THBS1</i>	0.766
14	<i>CAV1</i>	<i>FHL2</i>	0.765
15	<i>CAV1</i>	<i>IGFBP7</i>	0.765
16	<i>CAV1</i>	<i>TAGLN</i>	0.765
17	<i>CAV1</i>	<i>THBS1</i>	0.765
18	<i>C1QTNF1</i>	<i>CLIC3</i>	0.762
19	<i>FHL2</i>	<i>IGFBP3</i>	0.758
20	<i>DPYSL5</i>	<i>INSM1</i>	0.758
21	<i>CLIC3</i>	<i>TRIM29</i>	0.756
22	<i>C1QTNF1</i>	<i>KRT80</i>	0.756
23	<i>FHL2</i>	<i>WISP2</i>	0.754
24	<i>THBS1</i>	<i>WISP2</i>	0.754
25	<i>CAV1</i>	<i>MRGPRF</i>	0.754
26	<i>ARHGDI B</i>	<i>IL7R</i>	0.754
27	<i>CXXC4</i>	<i>DPYSL5</i>	0.752
28	<i>ABCC3</i>	<i>IER3</i>	0.752
29	<i>CLDN4</i>	<i>TACSTD2</i>	0.751
30	<i>CYR61</i>	<i>MFAP5</i>	0.751

↑ cut-off 0.75

Supplementary Table 2 | Combined *PC-corr* values calculated as minimum values of the two analyzed sets specifying the values for network edges. Edges above a cut-off of 0.70 are displayed. Cut-offs 0.70 and 0.75 are marked with a horizontal line.

	node <i>i</i>	node <i>j</i>	edge $PC-corr_{i,j}^{comb}$	
1	<i>FHL2</i>	<i>THBS1</i>	0.803	
2	<i>FHL2</i>	<i>IGFBP7</i>	0.785	
3	<i>IGFBP7</i>	<i>THBS1</i>	0.785	
4	<i>FHL2</i>	<i>TAGLN</i>	0.782	
5	<i>IGFBP7</i>	<i>TAGLN</i>	0.782	
6	<i>TAGLN</i>	<i>THBS1</i>	0.782	
7	<i>ANKRD1</i>	<i>IL11</i>	0.781	
8	<i>CAV1</i>	<i>FHL2</i>	0.759	
9	<i>CAV1</i>	<i>IGFBP7</i>	0.759	
10	<i>CAV1</i>	<i>TAGLN</i>	0.759	
11	<i>CAV1</i>	<i>THBS1</i>	0.759	
12	<i>CLIC3</i>	<i>TRIM29</i>	0.751	↑ cut-off 0.75
13	<i>ATP8B1</i>	<i>FHL2</i>	0.748	
14	<i>CAV1</i>	<i>MRGPRF</i>	0.736	
15	<i>FHL2</i>	<i>IGFBP3</i>	0.733	
16	<i>ARHGDIB</i>	<i>IL7R</i>	0.726	
17	<i>CYR61</i>	<i>MFAP5</i>	0.725	
18	<i>IGFBP3</i>	<i>TAGLN</i>	0.718	
19	<i>MAL</i>	<i>TRIM29</i>	0.718	
20	<i>FHL2</i>	<i>LBH</i>	0.717	
21	<i>FHL2</i>	<i>MRGPRF</i>	0.716	
22	<i>MRGPRF</i>	<i>THBS1</i>	0.709	
23	<i>CLIC3</i>	<i>SYT8</i>	0.709	
24	<i>LBH</i>	<i>TAGLN</i>	0.706	
25	<i>IGFBP6</i>	<i>MRGPRF</i>	0.705	
26	<i>ASCL1</i>	<i>LRRN2</i>	0.703	
27	<i>ATP8B1</i>	<i>MRGPRF</i>	0.702	
28	<i>SYT8</i>	<i>TRIM29</i>	0.701	
29	<i>CYR61</i>	<i>THBS1</i>	0.700	↑ cut-off 0.70

Supplementary Table 3 | List of target genes together with their processed PC loadings.

	gene name	gene description	V_t^1	V_t^2	\bar{V}
1	<i>ABCC3</i>	ATP binding cassette subfamily C member 3	0.849	0.693	0.771
2	<i>ANKRD1</i>	ankyrin repeat domain 1	0.933	0.781	0.857
3	<i>ARHGDI3</i>	Rho GDP dissociation inhibitor beta	0.726	0.933	0.829
4	<i>ASCL1</i>	achaete-scute family bHLH transcription factor 1	-0.703	-0.813	-0.758
5	<i>ATP8B1</i>	ATPase phospholipid transporting 8B1	0.813	0.748	0.780
6	<i>C1QTNF1</i>	C1q and TNF related 1	0.895	0.697	0.796
7	<i>CAV1</i>	caveolin 1	0.772	0.759	0.765
8	<i>CLDN4</i>	claudin 4	0.754	0.919	0.836
9	<i>CLIC3</i>	chloride intracellular channel 3	0.827	0.794	0.810
10	<i>CNN2</i>	calponin 2	0.673	0.920	0.796
11	<i>CXXC4</i>	CXXC finger protein 4	-0.614	-0.891	-0.752
12	<i>CYR61</i>	cellular communication network factor 1	0.777	0.754	0.765
13	<i>DPYSL5</i>	dihydropyrimidinase like 5	-0.686	-0.970	-0.828
14	<i>FHL2</i>	four and a half LIM domains 2	0.951	0.927	0.939
15	<i>IER3</i>	immediate early response 3	0.841	0.918	0.879
16	<i>IGFBP3</i>	insulin like growth factor binding protein 3	0.733	0.904	0.819
17	<i>IGFBP6</i>	insulin like growth factor binding protein 6	0.879	0.749	0.814
18	<i>IGFBP7</i>	insulin like growth factor binding protein 7	0.790	0.785	0.788
19	<i>IL11</i>	interleukin 11	0.845	0.880	0.862
20	<i>IL7R</i>	interleukin 7 receptor	0.788	-0.786	0.001
21	<i>INSM1</i>	INSM transcriptional repressor 1	-0.739	-0.979	-0.859
22	<i>KRT80</i>	keratin 80	0.872	0.639	0.756
23	<i>LBH</i>	LBH regulator of WNT signaling pathway	0.731	0.717	0.724
24	<i>LRRC15</i>	leucine rich repeat containing 15	0.916	-0.617	0.149
25	<i>LRRN2</i>	leucine rich repeat neuronal 2	-0.721	-0.733	-0.727
26	<i>MAL</i>	mal, T cell differentiation protein	0.724	0.841	0.782
27	<i>MFAP5</i>	microfibril associated protein 5	0.921	0.725	0.823
28	<i>MRGPRF</i>	MAS related GPR family member F	0.775	0.751	0.763
29	<i>SYT8</i>	synaptotagmin 8	0.746	0.709	0.727
30	<i>TACSTD2</i>	tumor associated calcium signal transducer 2	0.689	0.833	0.761
31	<i>TAGLN</i>	transgelin	0.782	0.782	0.782
32	<i>THBS1</i>	thrombospondin 1	0.922	0.803	0.863
33	<i>TRIM29</i>	tripartite motif containing 29	0.751	0.760	0.756
34	<i>WISP2</i>	cellular communication network factor 5	0.834	0.674	0.754

Supplementary Table 4 | Carcinoma cell lines. List of all carcinoma cell lines acquired from RIKEN institute used in this study, together with the catalogue number, tissue of origin, carcinoma type, growth medium specification, and passage number at purchase. non-sc – non small-cell, sq – squamous cell, ad – adenocarcinoma.

cell line	cat no	tissue	type	medium (Gibco cat no)	serum	passage
ECC4	RCB0982	intestine	small-cell	RPMI1640 (11875093)	10%	7
TGBC18TKB	RCB1169	intestine	non-sc (ad)	DMEM (11885084)	5%	5
WA-hT	RCB2279	lung	small-cell	MEM (11095080)	10%	54
EBC-1	RCB1965	lung	non-sc (sq)	MEM (11095080)	10%	7
A549	RCB0098	lung	non-sc (ad)	DMEM (11885084)	10%	92
ECC10	RCB0983	stomach	small-cell	RPMI1640 (11875093)	10%	8
MKN45	RCB1001	stomach	non-sc (ad)	RPMI1640 (11875093)	10%	6
MKN1	RCB1003	stomach	non-sc (ad)	RPMI1640 (11875093)	10%	6

Supplementary Table 5 | Mechanical characterizations of cells from the individual datasets using RT-DC. For each dataset experimental details of the measuring conditions are listed, including the widths of channel constriction ($w_{channel}$), total flow rates (Q_{total}), percentages of methylcellulose (MC) in the measurement buffer (buffer % MC), effective viscosity of the measurement buffer in the channel at the flowrate used (η_{eff} , according to ref.¹), as well as gates used for data filtering.

	measurement conditions				data filtering	
	$w_{channel}$ (μm)	Q_{total} ($\mu\text{l s}^{-1}$)	buffer % MC	η_{eff} (mPa s^{-1})	area (μm^2)	area ratio
glioblastoma	30	0.16	0.5	5.4	50–600	1.0–1.05
carcinoma	30	0.16	0.5	5.4	60–600	1.0–1.05
HSPCs	20	0.06	0.6	6.3	50–175	1.0–1.08
MCF10A	20	0.04	0.5	5.7	75–320	1.0–1.05
iPSCs	20	0.04	0.5	5.7	50–500	1.0–1.05
dev neurons	20	0.04	0.5	5.7	25–300	1.0–1.05
MEFs	30	0.16	0.5	5.4	50–500	1.0–1.05

Supplementary Table 6 | siRNAs used in the knock-down experiments.

name	target	commercial name	cat no	vendor
rLuc	Renilla Luciferase	RLUC	RLUC	Eupheria Biotec
esiCAV1-1	human caveolin 1	hCAV1	HU-03125-1	Eupheria Biotec
esiCAV1-2	human caveolin 1	hCAV1, custom design	HU-03125-2	Eupheria Biotec
esiCAV1-3	human caveolin 1	hCAV1, custom design	HU-03125-3	Eupheria Biotec
nonT	non-targeting	ON-TARGETplus Non-targeting Pool	D-001810-10-05	Dharmacon
CAV1-pool	human caveolin 1	ON-TARGETplus Human CAV1 siRNA, SMARTPool	L-003467-00-0005	Dharmacon

Supplementary Table 7 | Summary of details regarding transcriptomic profiling of the datasets used in this study.

	glioblastoma	carcinoma	HSPCs	MCF10A	iPSCs	developing neurons
accession no	GEO: GSE77751	DDBJ: DRA000991	GEO: GSE90552	GEO: GSE69822	GEO: GSE49940	GEO: GSE51606
technology	HT seq	CAGE	HT seq	HT seq	microarray	HT seq
instrument	Illumina HiSeq 2500	Helicos HeliScope	Illumina HiSeq 2500	Illumina HiSeq 2500	Illumina BeadArray Reader*	Illumina HiSeq 2000
platform ID	GPL16791	GPL14761	GPL16791	GPL16791	GPL6885	GPL13112
sequencing depth	26–35 million single-end reads per sample	4 million mapped tags per sample	27–56 million fragments per pooled libraries	31 million single-end reads per library	N/A	30–40 million reads per sample
RNA isolation	total RNA, High Pure RNA Isolation Kit (Roche)	total RNA, miRNeasy Kit (Qiagen)	total RNA, trizol isolation	total RNA, RNeasy Kit (Qiagen)	total RNA, RNeasy Kit (Qiagen)	polyA RNA, μ MACS™ mRNA Isolation Kit (Miltenyi)
library preparation	Ultra Directional RNA Library Prep (NEB)	HeliScopeCAGE ²	TruSeq RNA Sample Prep Kit (Illumina)	TruSeq RNA Sample Prep Kit (Illumina)	labelling with biotin	custom protocol
alignment	to GRCh38 GSNAP (v 2014-12-17)	to GRCh37 Delve	to GRCh38 GSNAP (v 2015-12-31)	to GRCh37 TopHat (v 2.0.10)	BeadStudio (v 3.2)	to MGSCv37 BWA (v 0.5.9)
counting	featureCounts (v 1.4.6)	decomposition peak identification (DPI) ^{3,\$}	featureCounts (v 1.5.0)	HTSeq (v 0.6.1) [†]	BeadStudio (v 3.2)	BEDtools (v 2.11)
normalization	size factor normalization (DESeq2 v 1.6.2)	TPM - tags per million (edgeR)	size factor normalization (DESeq2 v 1.10.1)	RPKM (DESeq2 v 1.4.5)	log2-scaling, qnt (Lumi)	size factor normalization (DESeq v 1.8.1)

*with mouseRef-8 v2 expression BeadChips (Illumina),

^{\$}available at <https://github.com/hkawaji/dpi1/>,

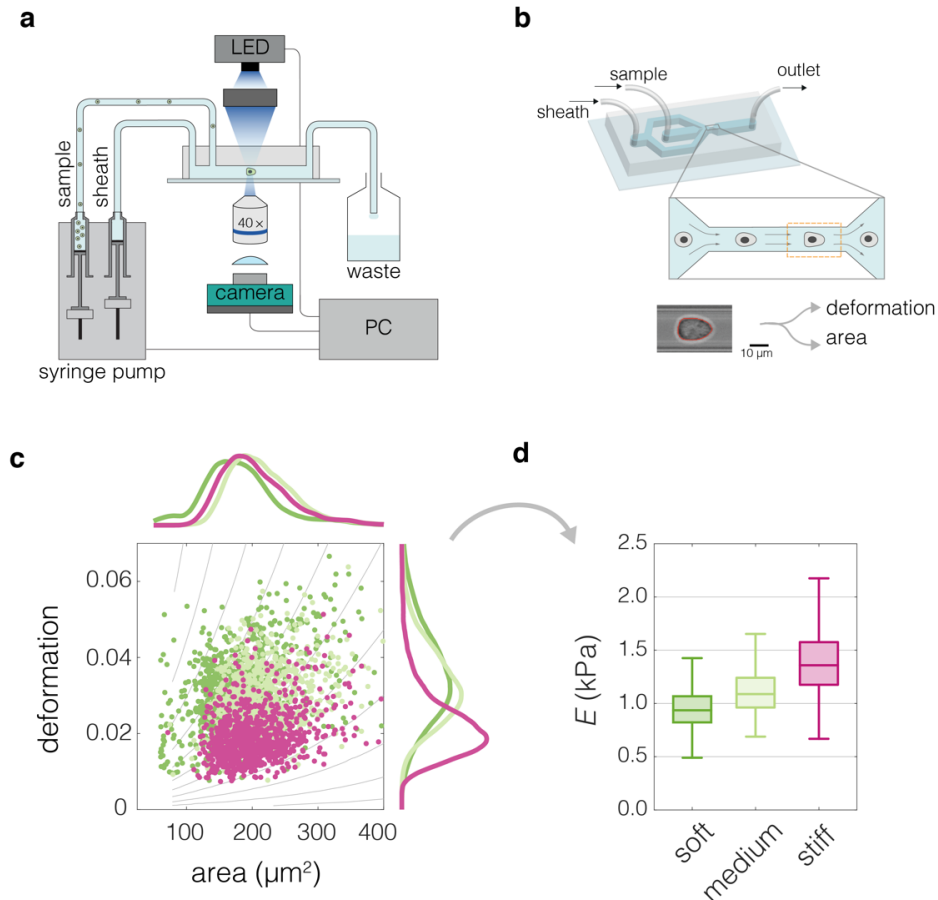
[†]available at <http://www-huber.embl.de/users/anders/HTSeq/doc/count.html>

Supplementary Table 8 | List of sample IDs assigned to the different cell states in the respective transcriptomic datasets. sc – small-cell, sq – squamous cell, adeno – adenocarcinoma.

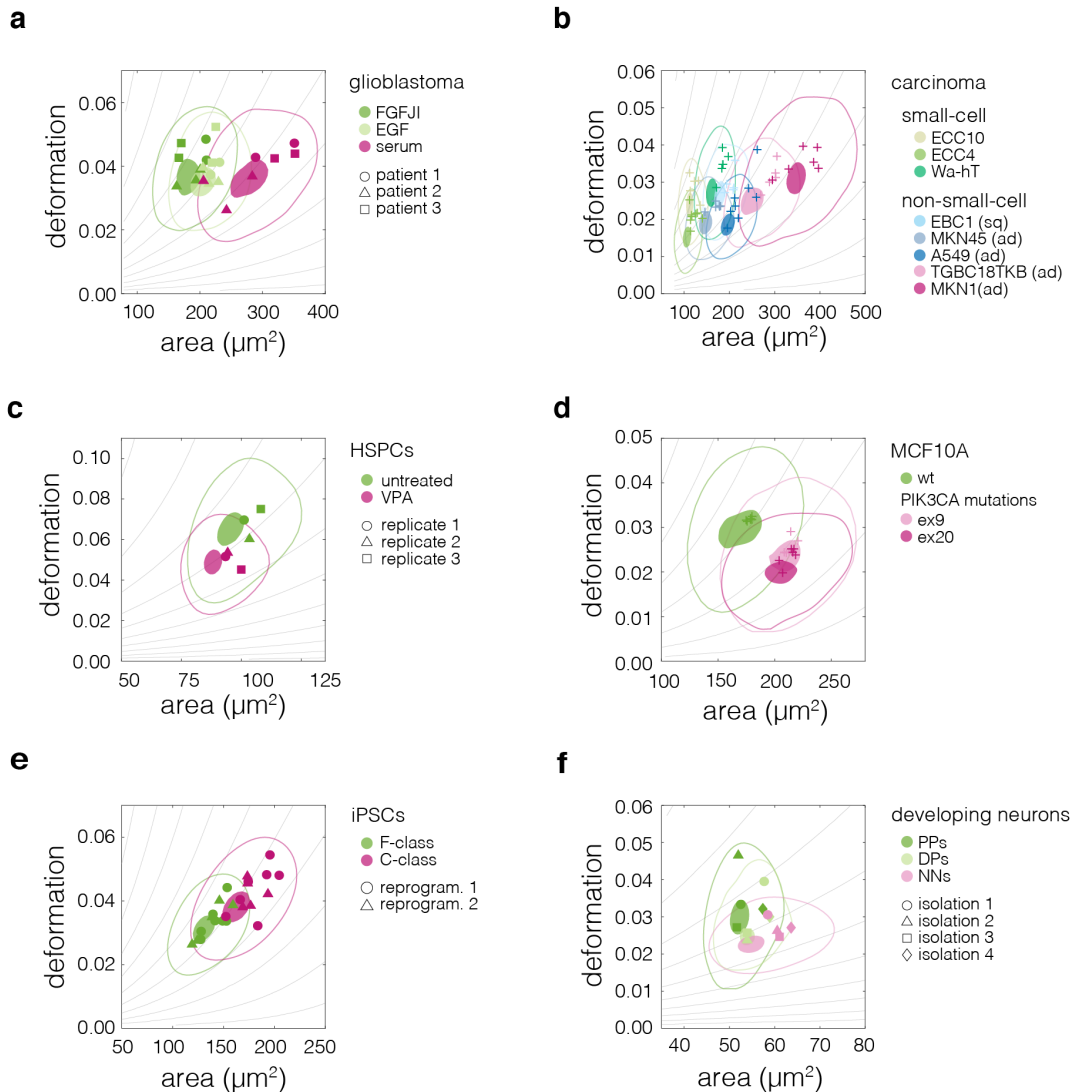
dataset accession No	cell state	sample IDs
glioblastoma GSE77751	FGFJl (soft)	GSM2058533 GSM2058534 GSM2058535 GSM2058542 GSM2058543 GSM2058544 GSM2058551 GSM2058552 GSM2058553
	EGF medium)	GSM2058530 GSM2058531 GSM2058532 GSM2058539 GSM2058540 GSM2058541 GSM2058548 GSM2058549 GSM2058550
	serum (stiff)	GSM2058536 GSM2058537 GSM2058538 GSM2058545 GSM2058546 GSM2058547 GSM2058554 GSM2058555 GSM2058556
carcinoma DRA000991	sc (soft)	10589 10610* 10841 10541 10842 10562 ^{\$} 10609 [§]
	sq (medium)	10717 10760 10692 10434 10550 10545 10544 10463 10486 ^{\$}
	adeno (stiff)	10796 10643 10614* 10612 10499 ^{\$} 10408 10648 10784 10437 10417 [§] 10639 11843 11841 10693 10797
HSPCs GSE90552	VPA (soft)	GSM2406738 GSM2406739 GSM2406740 GSM2406741
	PBS (stiff)	GSM2406734 GSM2406735 GSM2406736 GSM2406737
MCF10A GSE69822	WT (soft)	GSM1709515 GSM1709516 GSM1709517
	H1047R (stiff)	GSM1709572 GSM1709573 GSM1709574
iPSCs GSE49940	F-class (soft)	GSM1544134 GSM1544135 GSM1544139 GSM1544140 GSM1544146 GSM1544160
	C-class (stiff)	GSM1544136 GSM1544137 GSM1544138 GSM1544141 GSM1544142 GSM1544143 GSM1544144 GSM1544145 GSM1544147 GSM1544148 GSM1544149 GSM1544150 GSM1544151 GSM1544152 GSM1544153 GSM1544154 GSM1544155 GSM1544156 GSM1544157 GSM1544158 GSM1544159 GSM1544161
developing neurons GSE51606	PPs (soft)	GSM1249110 GSM1249113 GSM1249116
	DPs (medium)	GSM1249111 GSM1249114 GSM1249117
	NNs (stiff)	GSM1249112 GSM1249115 GSM1249118

*stomach, ^{\$}lung, [§]intestine

Supplementary Figures

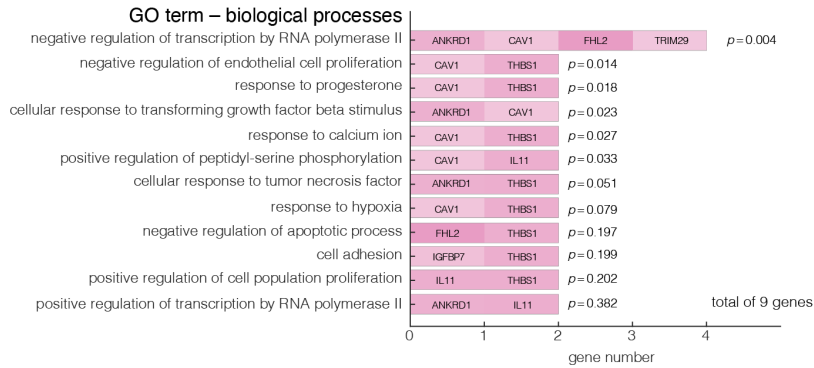


Supplementary Figure 1 | Characterization of mechanical cell properties using real-time deformability cytometry (RT-DC). **a**, Schematic overview of the RT-DC setup. Computer-operated syringe pumps flow the cell-containing sample as well as the sheath fluid into the microfluidic chip. Imaging of the cells deformed in the microfluidic channel is performed at 2,000 frames per second using an LED-based stroboscopic illumination and a CMOS camera. **b**, 3D illustration of the microfluidic chip used for the RT-DC measurements, close-up depicts the constriction of the channel in which cells are deformed, the imaged region of interest is indicated by an orange dashed line. At the bottom an exemplary image of a cell is shown. A contour is fitted to the cell in real time (marked in red), based on which cell area and deformation are calculated. **c**, An exemplary plot of deformation versus area of three different cell populations. The gray isoelasticity lines in the background indicate regions of the same apparent Young's moduli. **d**, Box plot of apparent Young's modulus, E , estimated based on deformation and area in **c**. The cell population with same area but higher deformation has lower E (bright green compared to magenta). For cells with similar deformation, the one of smaller area has lower E (dark green compared to bright green). The exemplary data in **c** and **d** corresponds to exemplary measurements of Wa-hT (dark green), EBC1 (bright green) and A549 (magenta) cell lines. The box plots in **d** spread from 25th to 75th percentiles with a line at the median, whiskers span 1.5x interquartile range (IQR).

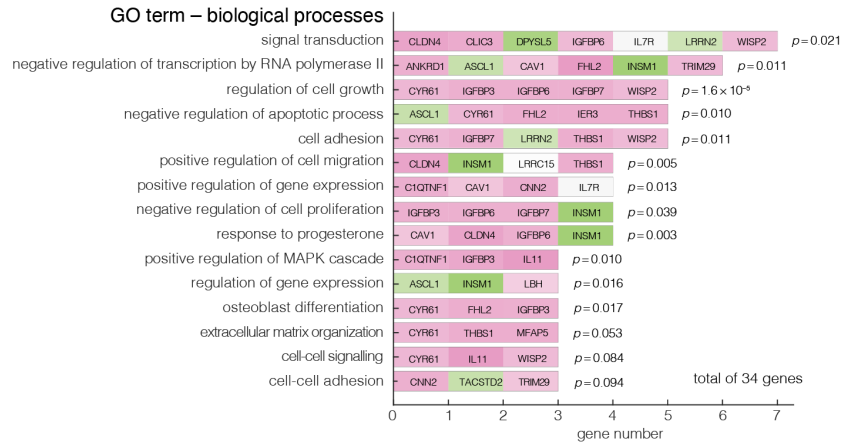


Supplementary Figure 2 | Plots of area vs deformation for different cell states in characterized systems. Panels correspond to the following systems: **a**, glioblastoma, **b**, carcinoma, **c**, human stem and progenitor cells (HSPCs), **d**, non-tumorigenic breast epithelia MCF10A, **e**, induced pluripotent stem cells (iPSCs), and **f**, developing neurons. 95%- and 50% density contours of data pooled from all measurements of given cell state are indicated by shaded areas and continuous lines, respectively. Datapoints indicate medians of individual measurements. The isoelasticity lines in the background (gray) indicate regions of the same apparent Young's moduli. DDs – differentiating progenitors, DPs – differentiating progenitors, NNs – newborn neurons.

a

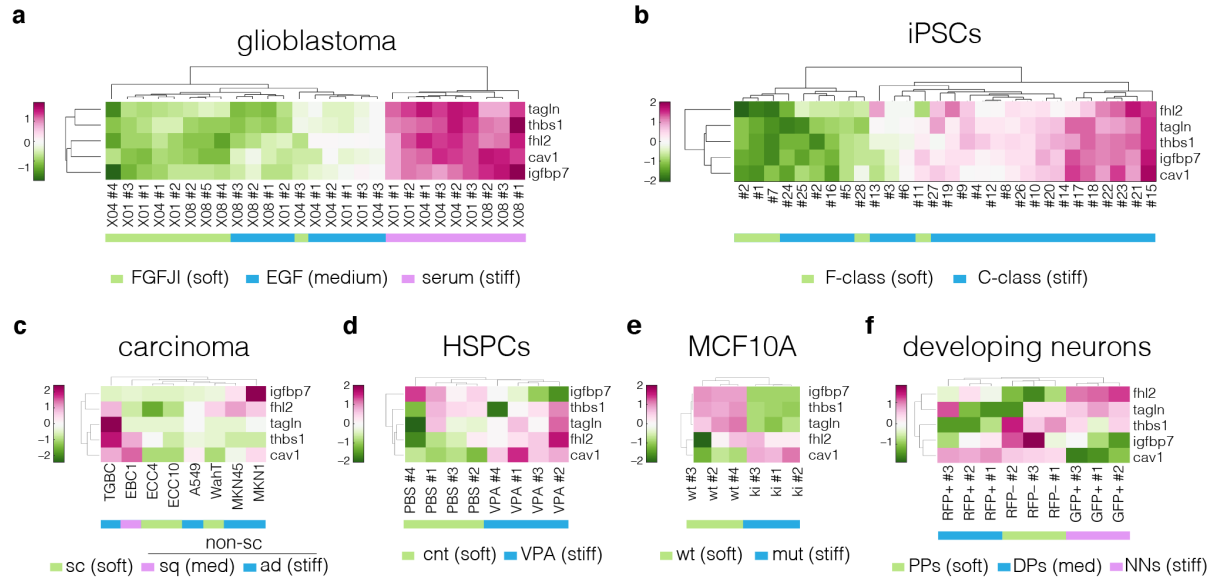


b

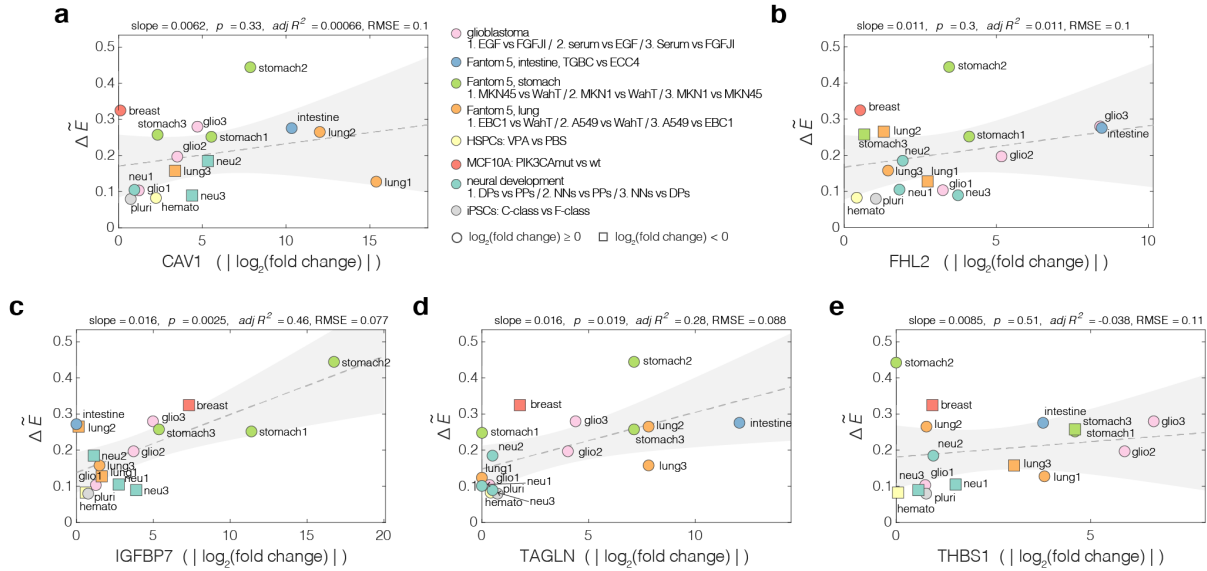


Supplementary Figure 3 | Gene ontology (GO) enrichment analysis of obtained target genes.

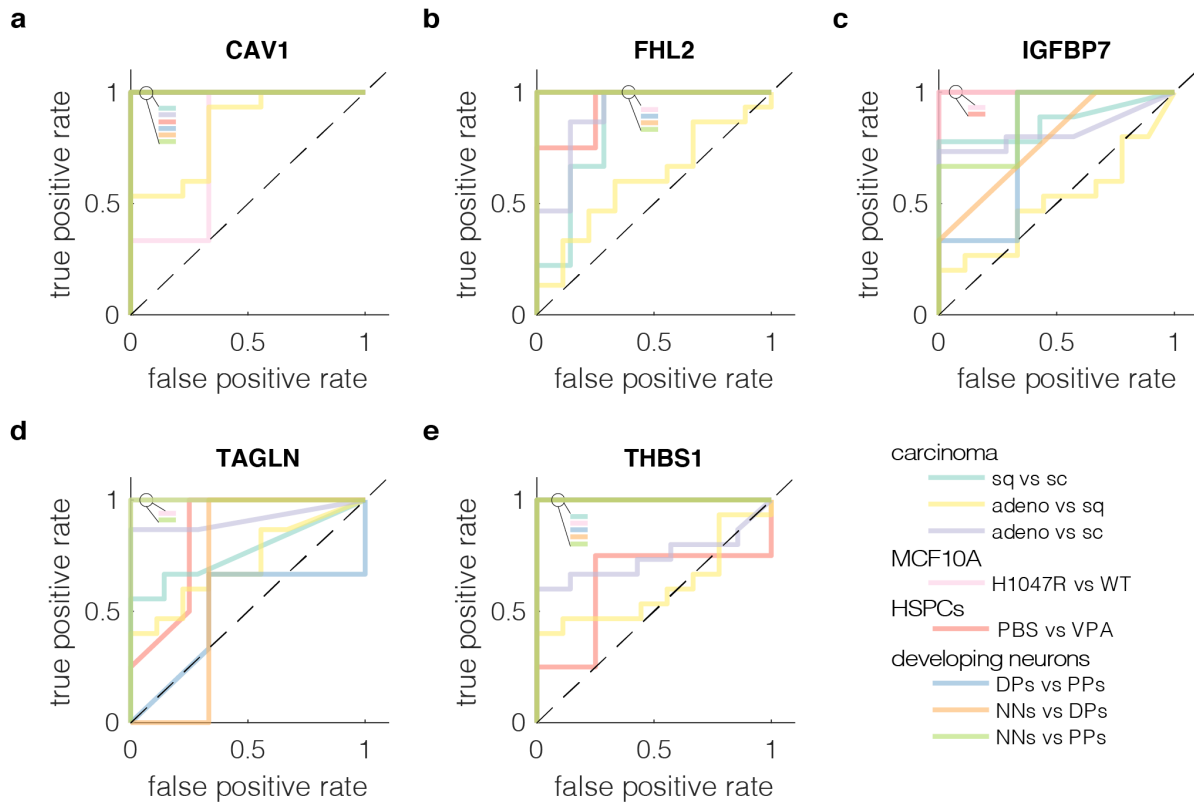
Enriched GO terms of biological processes are summarized for **a**, 9 genes corresponding to the results from Figure 3g, **b**, 34 genes corresponding to all nodes presented in Figure 3e–g. The analysis was performed using *DAVID 6.8* functional annotation tool online, with *Homo sapiens* as background dataset, ENSEMBL gene IDs as input, and focused on direct GO terms for biological processes. Color code of the blocks corresponds to the level of expression in stiff states with green corresponding to low expression and magenta corresponding to high expression. The reported p -values are the Fisher's exact p -values obtained using a two tailed two sample t-test.



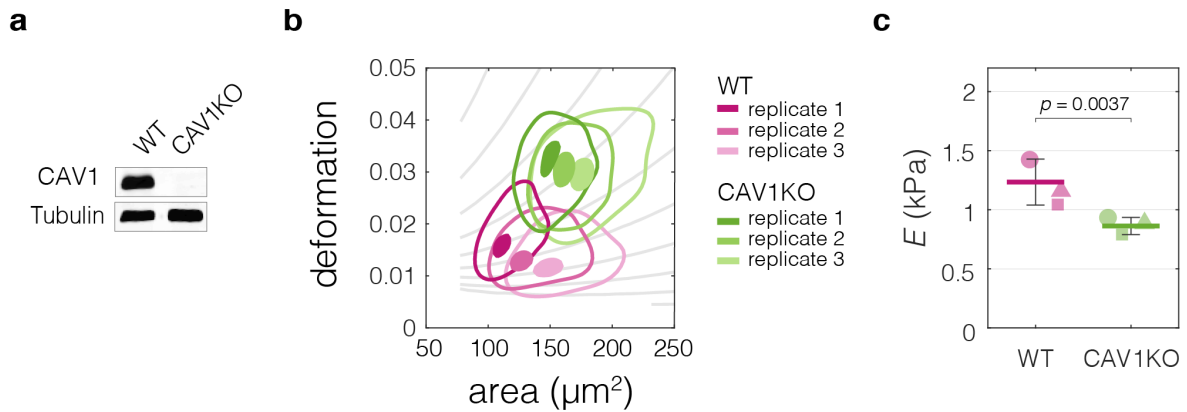
Supplementary Figure 4 | Expression of identified target genes in the discovery and validation datasets. Panels show unsupervised clustering heat maps of expression data from transcriptomic datasets corresponding to the following systems: **a**, glioblastoma, **b**, carcinoma, **c**, human stem and progenitor cells (HSPCs), **d**, non-tumorigenic breast epithelia MCF10A, **e**, induced pluripotent stem cells (iPSCs), and **f**, developing neurons. sc – small-cell carcinoma, non-sc – non-small-cell carcinoma, sq – squamous cell carcinoma, adeno – adenocarcinoma, cnt – untreated control, VPA – valproic acid, wt – wild type, mut – mutant, PPs – proliferating progenitors, DPs – differentiating progenitors, NNs – newborn neurons. Clustering was performed using *clustergram* function in *MATLAB (R2020a, MathWorks)* on log2 of expression data.



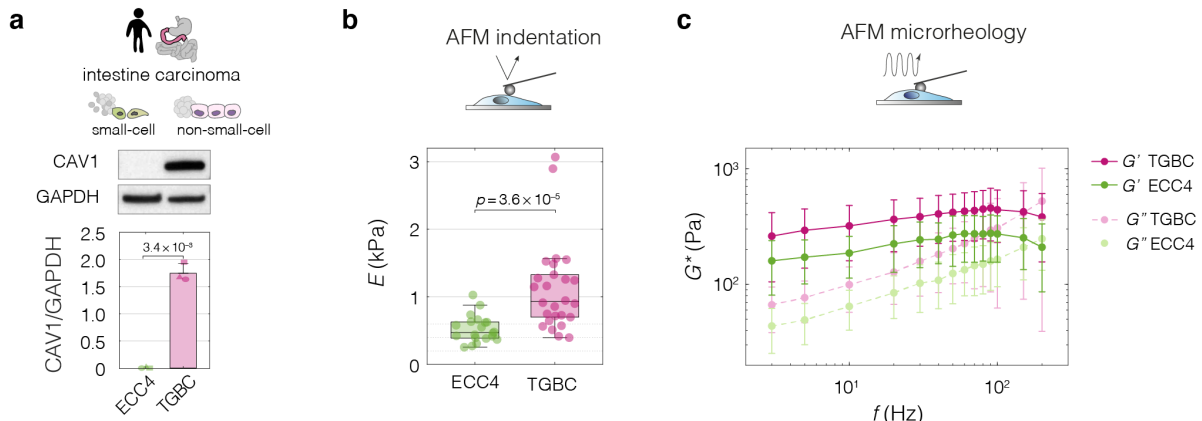
Supplementary Figure 5 | Relation between the magnitude of apparent Young's modulus change and the absolute change in the expression levels of target genes. Plots of normalized change in apparent Young's modulus $\Delta\tilde{E}$ versus absolute value of change in expression for the target genes from conserved module: **a**, CAV1, **b**, FHL2, **c**, IGFBP7, **d**, TAGLN, **e**, THBS1. Every soft-stiff state pair from the respective datasets is presented as an individual point. $\Delta\tilde{E} = \frac{E_{stiff} - E_{soft}}{E_{stiff}}$, where E_{stiff} and E_{soft} correspond to the apparent Young's moduli (mean of all measurements) of the stiff and soft states within the given pairs, respectively. The dashed lines correspond to linear fits to data, with gray-shaded areas representing 95% confidence intervals.



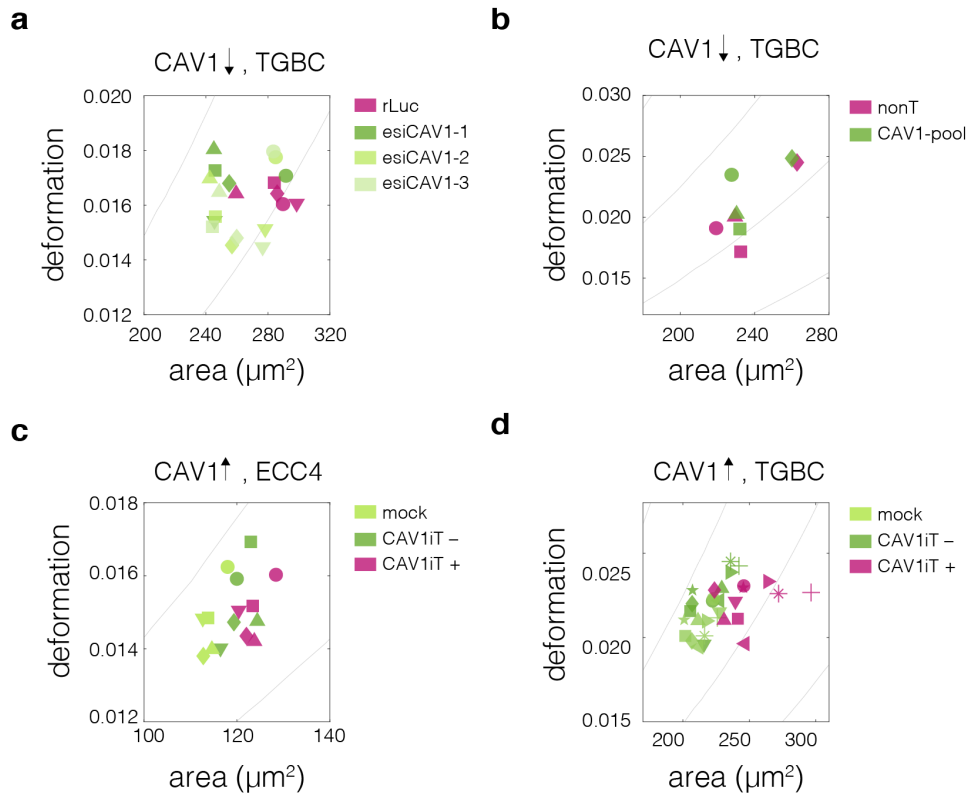
Supplementary Figure 6 | ROC curves characterizing classification performance of the five genes from the conserved module. True positive rate was plotted against the false positive rate at different classification thresholds for each soft-stiff phenotype pair from the validation datasets for: **a**, CAV1, **b**, FHL2, **c**, IGFBP7, **d**, TAGLN, and **e**, THBS1. The insets in the upper left corners of the plot show the colors of all overlying curves with AUC = 1. The ROC curves were constructed using *perfcurve* function in *MATLAB (R2020a, MathWorks)*. sc – small cell carcinoma, sq – squamous cell carcinoma, adeno – adenocarcinoma, wt, wild type, cnt – control, VPA – valproic acid, PPs – proliferating progenitors, DPs – differentiating progenitors, NNs – newborn neurons.



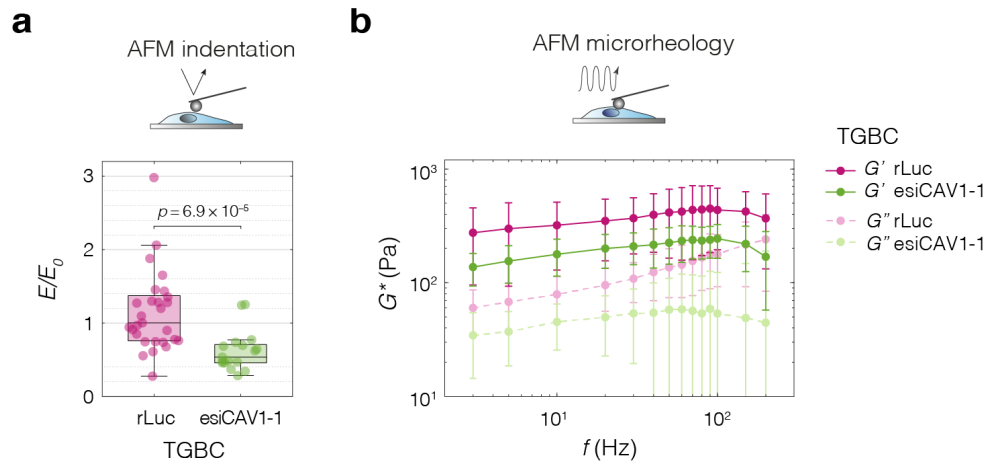
Supplementary Figure 7 | CAV1 knock-out mouse embryonic fibroblasts (CAV1KO) have lower stiffness compared to the wild type cells (WT). **a**, Western blot analysis of CAV1 expression levels in CAV1KO compared to WT cells. **b**, Plots of area vs deformation for CAV1KO and WT cells characterized with RT-DC. Contour plots delineate 95% and 50% density areas (solid lines and filled area, respectively) of data from individual measurement replicates ($n = 3$). The isoelasticity lines in the background (gray) indicate regions of the same apparent Young's moduli. **c**, Apparent Young's modulus values estimated for WT and CAV1KO cells using area-deformation data in **a**. Horizontal lines delineate medians with mean absolute deviation (MAD) as error, datapoints represent medians of the individual replicates. Statistical analysis was performed using generalized linear mixed effects model.



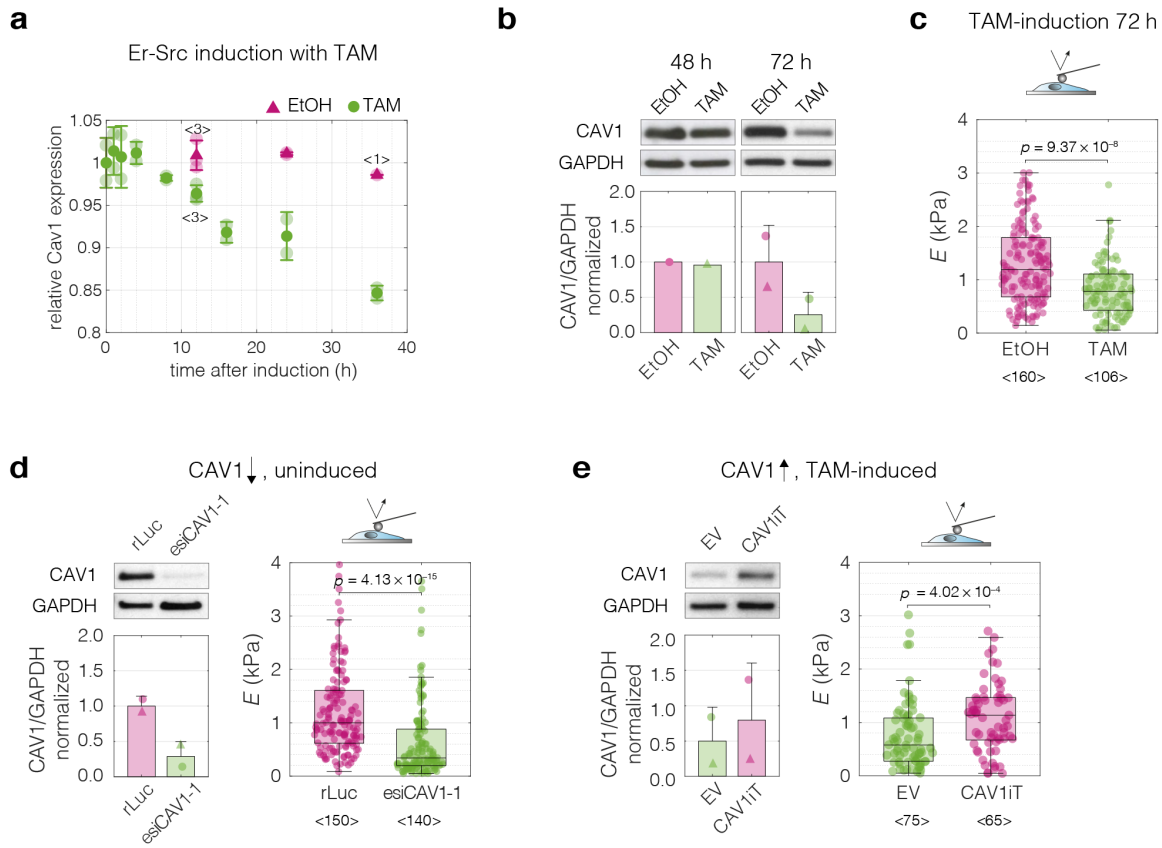
Supplementary Figure 8 | CAV1 expression and mechanical characterization with AFM of small-cell (ECC4) vs non-small-cell (TGBC) carcinoma cell lines from intestine. **a**, ECC4 do not show detectable levels of CAV1, while TGBC have considerable basal CAV1 expression. For the Western blot analysis representative blots (top) as well as quantification (bottom, $n = 3$) are shown. **b**, ECC4 show lower apparent Young's moduli than TGBC in AFM indentation experiments. Box plots spread from 25th to 75th percentiles with a line at the median, whiskers span 1.5x interquartile range (IQR), individual datapoints correspond to values obtained for individual cells (number of measured cells $n = 20$ and 26 for ECC4 and TGBC, respectively). Statistical analysis was performed using two-sided two-sample t-test (a) or two-sided Wilcoxon rank sum test (b). **c**, ECC4 show storage and shear moduli lower than TGBC in AFM microrheology measurements. Datapoints correspond to means \pm standard deviation of all measurements at given oscillation frequencies ($n = 18$ and 27 for each frequency for ECC4 and TGBC, respectively). Lines connecting datapoints serve as guides for the eye.



Supplementary Figure 9 | Plots of area vs deformation from RT-DC measurements of cells with perturbed CAV1 levels. Panels correspond to the following experiments: **a–b**, CAV1 knock-down in TGBC cells using esiRNA (a) and ONTarget siRNA (b), **c–d**, transient CAV1 overexpression in ECC4 cells (c) and TGBC cells (d). Datapoints indicate medians of individual measurement replicates. The isoelasticity lines in the background (gray) indicate regions of same mechanical properties.



Supplementary Figure 10 | TGBC cells show decreased stiffness upon CAV1 knock-down as measured by AFM. a, After CAV1 knock-down (esiCAV1-1), the TGBC cells show lower apparent Young's moduli than control cells transfected with non-targeting esiRNA (rLuc). Box plots spread from 25th to 75th percentiles with a line at the median, whiskers span 1.5x interquartile range (IQR), individual datapoints correspond to values obtained for individual cells (number of measured cells $n = 29$ and 17 for rLuc and esiCAV1-1 conditions, respectively). Statistical analysis was performed using two sample two-sided Wilcoxon rank sum test. **b**, After CAV1 knock-down (esiCAV1-1), the TGBC cells show storage and shear modulus lower than the control cells (rLuc) in AFM microrheology measurements. Datapoints correspond to means \pm standard deviation of all measurements at given oscillation frequencies ($n = 32$ and 18 for each frequency for rLuc and esiCAV1-1 conditions, respectively). Lines connecting datapoints serve as guides for the eye.



Supplementary Figure 11 | Perturbations of CAV1 levels in MCF10A-ER-Src cells result in cell stiffness changes. **a**, Inducing transformation of MCF10A- ER-Src cells by tamoxifen (TAM) treatment, as opposed to vehicle control (ethanol, EtOH), causes a decrease of CAV1 expression over time, as captured by microarray analysis⁴. Datapoints with error bars represent means \pm standard deviation ($n = 2$, unless indicated otherwise). **b**, Western blot analysis shows the decrease of CAV1 at protein level 72 h post induction. **c**, MCF10A-ER-Src cells show decreased apparent Young's moduli 72 h post TAM induction. **d**, CAV1 knock-down in uninduced MCF10A-ER-Src cells results in lowering of the apparent Young's modulus. **e**, Overexpression of CAV1 in TAM-induced MCF10A-ER-Src cells causes increase in the apparent Young's modulus and effectively reverts the softening caused by TAM induction (compare to panel c). Box plots in c–e spread from 25th to 75th percentiles with a line at the median, whiskers span 1.5x interquartile range (IQR), individual datapoints correspond to values obtained for individual cells, the number of measured cells per conditions, pooled from $n = 3$ independent experiments, is indicated below each box. Statistical analysis was performed using a two-sided Wilcoxon rank sum test.

Supplementary References

1. Herold, C. Mapping of Deformation to Apparent Young's Modulus in Real-Time Deformability Cytometry. (2017).
2. Itoh, M. *et al.* Automated workflow for preparation of cDNA for cap analysis of gene expression on a single molecule sequencer. *PLoS One* **7**, (2012).
3. Forrest, A. R. R. *et al.* A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).
4. Hirsch, H. A. *et al.* A Transcriptional Signature and Common Gene Networks Link Cancer with Lipid Metabolism and Diverse Human Diseases. *Cancer Cell* **17**, 348–361 (2010).