De novo identification of universal cell mechanics gene signatures

Marta Urbanska[1,2,‡,†], Yan Ge[1,§,†], Maria Winzi[1], Shada Abuhattum[1,2], Syed Shafat Ali[3,4], Maik Herbig[1,2,5], Martin Kräter[1,2], Nicole Toepfner[1,6], Joanne Durgan[7], Oliver Florey[7], Martina Dori[5], Federico Calegari[5], Fidel-Nicolás Lolo[8], Miguel Á. del Pozo[8], Anna Taubenberger[1], Carlo V. Cannistraci[4,1,9,10,*], Jochen Guck[1,2,*]

[1]Biotechnology Center, Center for Molecular and Cellular Bioengineering, Technische Universität Dresden, Tatzberg 47-49, Dresden 01307, Germany

[2]Max Planck Institute for the Science of Light & Max-Planck-Zentrum für Physik und Medizin, Staudtstraße 2, Erlangen 91058, Germany

[3]Department of Computer Science and Department of Economics, Jamia Millia Islamia, New Delhi 110025, India

[4]Center for Complex Network Intelligence, Tsinghua Laboratory of Brain and Intelligence, Department of Computer Science and Department of Biomedical Engineering, Tsinghua University, Chengfu Rd. 160, SanCaiTang Building, Haidian District, Beijing 100084, China

[5]Center for Regenerative Therapies Dresden, Center for Molecular and Cellular Bioengineering, Technische Universität Dresden, Fetscherstrasse 105, Dresden 01307, Germany

[6]Klinik und Poliklinik für Kinder- und Jugendmedizin, Universitätsklinikum Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, Dresden 01307, Germany

[7]Signalling Programme, The Babraham Institute, Babraham Research Campus, Cambridge CB22 3AT, UK

[8]Mechanoadaptation and Caveolae Biology lab, Cell and Developmental Biology Area, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Calle de Melchor Fernández Almagro 3, Madrid 28029, Spain

[9]Center for Systems Biology Dresden, Pfotenhauerstraße 108, Dresden 01307, Germany

[10]Cluster of Excellence Physics of Life, Technische Universität Dresden, Tatzberg 47-49, Dresden 01307, Germany

[‡]Present address: Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3EL, UK

[§]Present address: AliveX Biotech, Shanghai 200030, China

[†]These authors contributed equally to this work

[*]Corresponding authors: kalokagathos.agon@gmail.com (C.V.C.), jochen.guck@mpl.mpg.de (J.G.)

Abstract

Cell mechanical properties determine many physiological functions, such as cell fate specification, migration, or circulation through vasculature. Identifying factors that govern the mechanical properties is therefore a subject of great interest. Here we present a mechanomics approach for establishing links between single-cell mechanical phenotype changes and the genes involved in driving them. We combine mechanical characterization of cells across a variety of mouse and human systems with machine learning-based discriminative network analysis of associated transcriptomic profiles to infer a conserved network module of five genes with putative roles in cell mechanics regulation. We validate in silico that the identified gene markers are universal, trustworthy and specific to the mechanical phenotype, and demonstrate experimentally that a selected target, *CAV1*, changes the mechanical phenotype of cells accordingly when silenced or overexpressed. Our data-driven approach paves the way towards engineering cell mechanical properties on demand to explore their impact on physiological and pathological cell functions.

Keywords: mechanical phenotype, deformability cytometry, transcriptomics, RNAseq, unsupervised machine learning, systems biology, discriminative network analysis, mechanomics

Introduction

The extent to which cells can be deformed by external loads is determined by their mechanical properties, such as cell stiffness. Since the mechanical phenotype of cells has been shown to reflect functional cell changes, it is now well established as a sensitive label-free biophysical marker of cell state in health and disease (*1*–*3*). Beyond being a passive property that can be correlated with cell state, cell stiffness is increasingly recognized as an important feature involved in processes such as development (*4*, *5*) and cancer progression (*6*, *7*). Identifying the molecular targets for on-demand tuning of mechanical properties is, thus, essential for exploring the precise impact that cell mechanics has on physiological and pathological processes in living organisms.

The mechanical properties of cells are determined by various intracellular structures and their dynamics, with cytoskeletal networks at the forefront (*8*). According to current knowledge, the most prominent contributor to the global mechanical phenotype is the actin cortex and its contractility regulated via Rho signaling (*9*, *10*). Intermediate filaments, including vimentin and keratin, reside deeper inside the cell and can also contribute to measured cell stiffness, especially at high strains (*11*, *12*). Although there is some evidence of the contribution of microtubules to cell stiffness at high strains (*13*), their role has been difficult to address directly, since drug-induced microtubule disassembly evokes reinforcement of actin cytoskeleton and cell contractility (*14*). Apart from cytoskeletal contributions, the cell mechanical phenotype can be influenced by the level of intracellular packing (*15*, *16*) or mechanical properties of organelles occupying the cell interior, such as the cell nucleus (*17*). When aiming at modulating the mechanical properties of cells, it may not be practical to target cytoskeletal structures, which are central to a multitude of cellular processes, because their disruption is generally toxic to cells. It is therefore important to identify targets that enable subtle, alternative ways of intervening with cell stiffness.

Most of our knowledge about the molecular contributors to cell mechanics has been derived from drug perturbations or genetic modifications targeting structures known a priori. The

challenge of identifying novel targets determining the mechanical phenotype can be addressed on a large scale by performing screens using RNA interference (*18–20*) or small-molecule compound libraries. Alternatively, the problem can be reverse-engineered, in that omics datasets for systems with known mechanical phenotype changes are used for prediction of genes involved in the regulation of mechanical phenotype in a mechanomics approach (*3*). Broadly speaking, mechanomics is a study of omics data within the context of mechanobiology. So far, this term has been used with regard to changes in omics profiles in response to an external mechanical stimulus such as shear flow, tensile stretch, or mechanical compression (*21–23*), or to collectively name all of the mechanical forces acting on or within cells (*24–28*). However, it can also be used to address omics changes related to changes in the mechanical properties of cells (*29*, *30*) — a context much closer to our study.

Here we extend the concept of mechanomics to a data-driven methodology for de novo identification of genes associated with the mechanical phenotype based on omics data (Fig. 1). To demonstrate this approach, we perform a machine learning-based discriminative network analysis termed PC-corr (*29*) on transcriptomics data from two unrelated biological systems with known mechanical phenotype changes (*30*, *31*) and elucidate a conserved functional module of five candidate genes putatively involved in the regulation of cell mechanics. We then test the ability of each gene to classify cell states according to cell stiffness in silico on six further transcriptomic datasets and show that the individual genes, as well as their compression into a combinatorial marker, are universally, specifically and trustworthily associated with the mechanical phenotype. Finally, we confirm experimentally that one of the candidate genes, caveolin-1 (*CAV1*), has the capacity to alter the mechanical phenotype in the predicted direction when downregulated or overexpressed. The systematic approach presented here, combining omics data with mechanical phenotypes across different systems, has the power to identify genes that ubiquitously contribute to cell mechanical phenotype in a hypothesis-free manner. Such genes can, in the future, be used as knobs for adjusting mechanical cell properties to explore their role in the homeostasis of multicellular systems or to therapeutically intervene in relevant pathologies.

Results

Cross-system identification of genes involved in cell mechanical changes

We introduce an inference approach for de novo identification of genes involved in cell mechanical changes across different systems that we refer to as mechanomics. The general workflow of this approach is presented in Fig. 1 and consists of four steps: data curation, prediction, in silico validation and experimental validation. In the first step, mechano-transcriptomic datasets representing a broad spectrum of biological systems are collected (Fig. 1A). Each dataset encompasses two or more cell states characterized by a distinct mechanical phenotype, and for which transcriptomic data is available. In the second step, a subset of the transcriptomic datasets is used to identify a conserved network module of putative target genes involved in the regulation of cell mechanical phenotype (Fig. 1B). The ability of the obtained target genes to correctly classify soft and stiff cell states is next tested in silico on the validation datasets (Fig. 1C). Finally, the best scoring targets are validated experimentally by monitoring mechanical phenotype changes upon their overexpression and downregulation in the cells of choice (Fig. 1D).

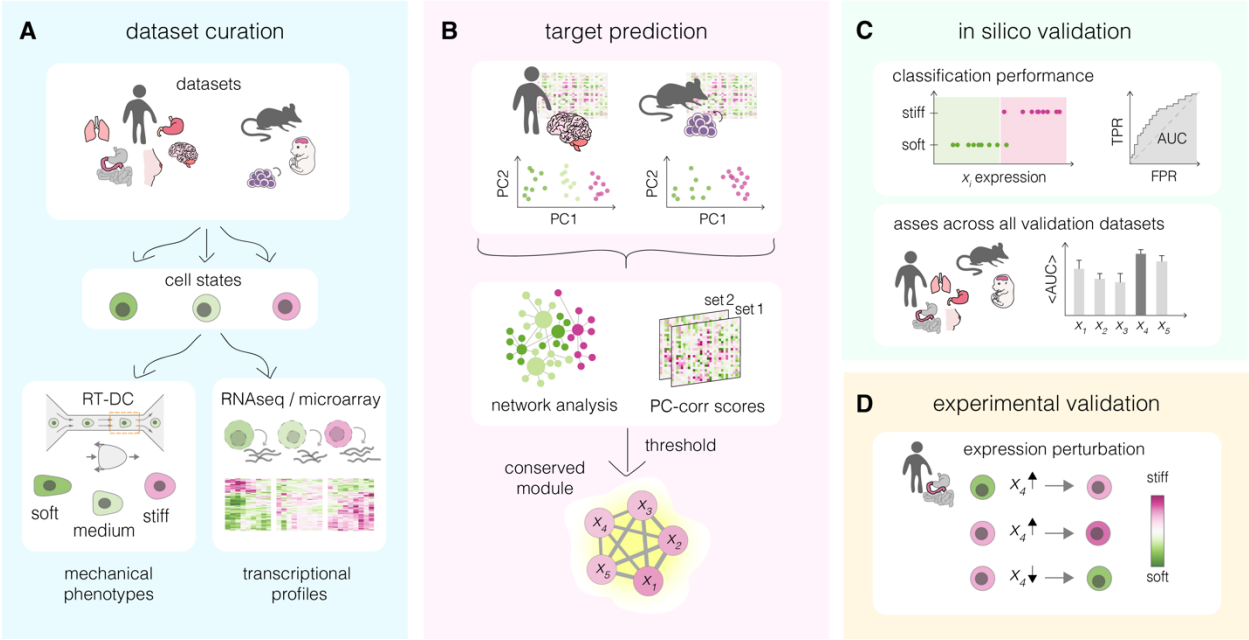

Fig. 1. Overview of a mechanomics approach for de novo identification of genes involved in cell mechanics regulation. (**A**) Data curation. Datasets originating from different biological systems encompassing cell states with distinct mechanical phenotypes, as characterized by real-time deformability cytometry (RT-DC), and associated transcriptomics profiles are collected. (**B**) Target prediction. A subset of collected datasets is used to perform machine learning-based network analysis on transcriptomic data and identify conserved module of genes associated with cell mechanics changes. PC – principal component. (**C**) In silico validation. The classification performance of individual genes from module identified in (B) is evaluated in silico on remaining datasets. TPR – true positive rate, FPR – false positive rate, AUC – area under the curve. (**D**) Experimental validation. Targets with highest classification performance in silico are verified experimentally in perturbation experiments.

Model systems characterized by mechanical phenotype changes

To curate the datasets, we screened the projects ongoing in our group and identified five biological systems for which published transcriptomic data were available, and the concomitant mechanical phenotype changes were either already documented or implicated (Table 1). The mechanical phenotypes of the different cell states within each dataset were characterized primarily using real-time deformability cytometry (RT-DC), a microfluidics-based method that enables rapid analysis of thousands of cells (*32*) (Fig. S1) — a feature particularly useful when setting out to explore a large variety of systems and states. RT-DC relies on flowing cells through a narrow constriction of a microfluidic channel and high-speed imaging to assess the ensuing cell deformation (*32*) (Fig. S1, A and B). In the context of this method, the mechanical phenotype is understood as whole-cell elasticity quantified by an apparent Young's modulus, *E*, deduced from cell size and deformation under given experimental conditions (*33*) (Fig. S1, C and D). Young's modulus quantifies how much stress (force per unit area) is necessary to deform a cell to a certain extent (i.e., strain), thus higher Young's modulus values indicate that a cell is more difficult to deform, or stiffer. In two of the datasets (see Table 1), selected cell states were additionally characterized using atomic force microscopy (AFM)-based assays on adherent cells to confirm the mechanical differences observed with RT-DC. The transcriptional profiles related to each system, generated by either RNA sequencing (RNA-Seq) or microarray analysis, were retrieved from entries previously deposited in online databases (Table 1).

We curated mechano-transcriptomic data assemblies originating from five different biological systems (Fig. 2) that included a total of eight transcriptomic datasets (Table 1). Two of the transcriptomic datasets were used for target prediction, and the reaming six for target validation. The first studied system encompassed patient-derived glioblastoma cell lines cultured in conditions supporting different levels of activation of the STAT3-Ser/Hes3 signaling axis involved in cancer growth regulation. As previously demonstrated, the higher the STAT3-Ser/Hes3 activation in the characterized states, the stiffer the measured phenotype of glioblastoma cells (*30*) (Fig. 2A).

Table 1. Mechano-transcriptomic datasets used in this study. Pred – prediction, Val – validation, PI/II– positive hypothesis I/II, N – negative hypothesis, CCLE – cancer cell line encyclopedia**,** HT Seq – high-throughput RNA sequencing, CAGE – cap analysis gene expression, AFM – atomic force microscopy, adeno – adenocarcinoma, wt – wild type, PP – proliferating progenitors, NNs – newborn neurons.

general information				transcriptomic data					mechanics data	
source	dataset name	used for	cell states	accession number	reference	method	unique entries	total samples used	method	reference
human	glioblastoma	Pred	FGFJI | EGF | serum	GEO: GSE77751	Poser *et al.* (*30*)	HT seq	39400	27	RT-DC	Poser *et al.* (*30*)
	carcinoma	Val PI	small-cell | adeno	DDBJ: DRA000991[§]	FANTOM5 consorptium (*34*)	CAGE	18821	12	RT-DC, AFM	this paper
		Val PII & N		GEO: GSE36139[†]	CCLE microarray Barretina *et al.* (*35*)	microarray	18925	162		
		Val PII		DepMap: release 21Q4[#]	CCLE RNA-Seq Ghandi *et al.* (*36*)	HT seq	51304	179		
		Val PII & N		GEO: GSE30611	Genentech Klijn *et al.* (*37*)	HT seq	25996	82		
	MCF10A	Val PI	wt | H1047R	GEO: GSE69822	Kiselev *et al.* (*38*)	HT seq	38508	6	RT-DC	this paper
mouse	iPSCs	Pred	F-class | C-class	GEO: GSE49940	Tonge *et al.* (*39*)	microarray	18118	28	RT-DC, AFM	Urbanska *et al.* (*31*)
	developing neurons	Val PI	PPs | NNs	GEO: GSE51606	Aprea *et al.* (*40*)	HT seq	21110	9	RT-DC	this paper

[§]data for samples of interest was extracted using TET tool from the FANTOM5 website https://fantom.gsc.riken.jp/5/

[†]data was downloaded using the ArrayExpress archive https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-2706/

[#]DepMap Public 21Q4 Primary Files, accessed via DepMap portal https://depmap.org/portal/downloa

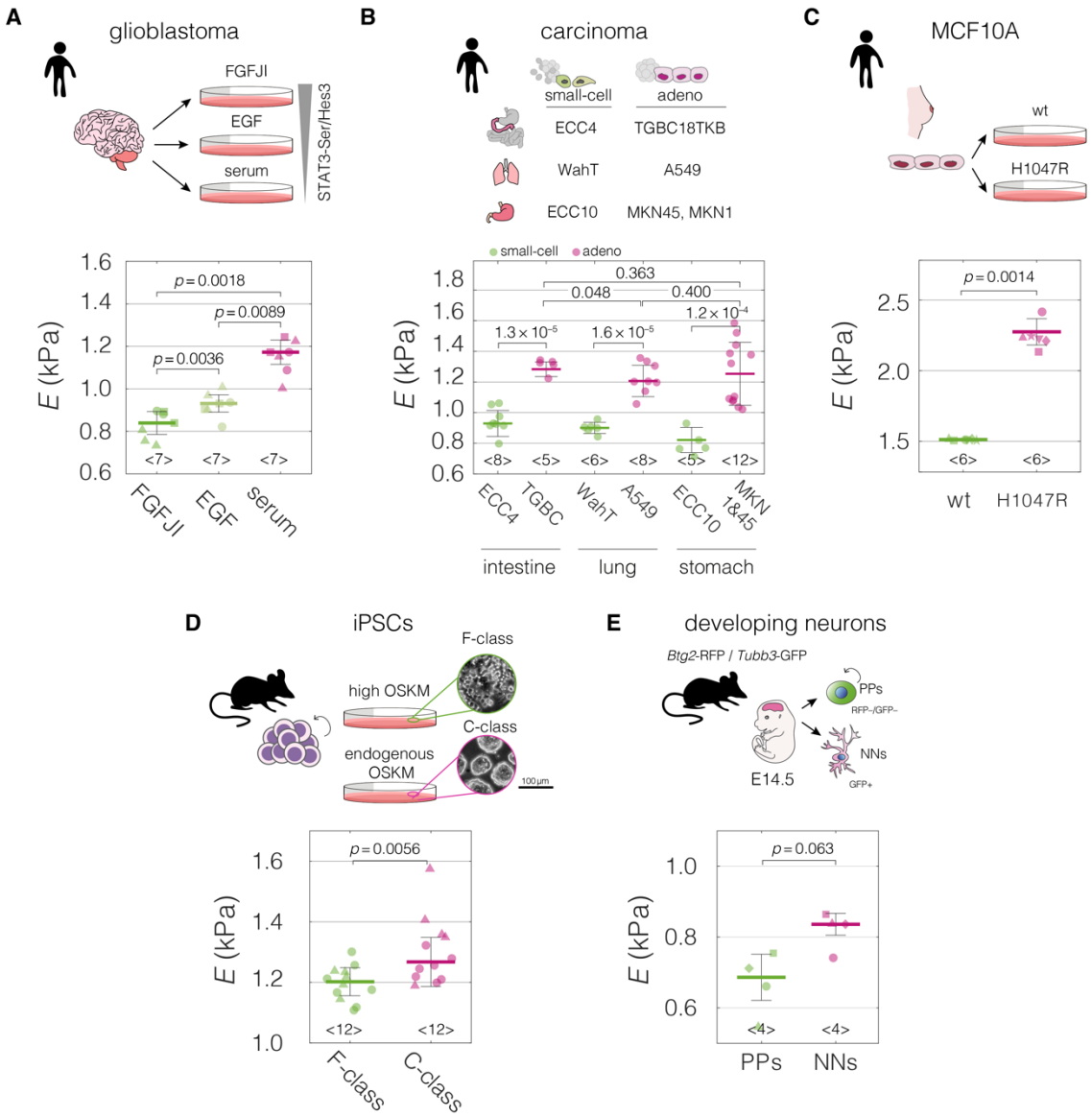

Fig. 2. Mechanical properties of divergent cell states in five biological systems. Schematic overviews of the systems used in our study, alongside with the cell stiffness of individual cell states parametrized by Young's moduli *E*. (**A**) Human patient-derived glioblastoma cells with three distinct signaling states maintained by indicated culture conditions. (**B**) Human small-cell carcinoma and adenocarcinoma cell lines originating from intestine, lung, and stomach. (**C**) Human breast epithelium MCF10A cell line bearing single-allele H1047R mutation in the PIK3CA with parental wild type (wt) as a control. (**D**) Murine F- and C-class iPSCs cultured in the presence or absence of doxycycline (dox) activating ectopic expression of OSKM factors (Oct4, Sox2, Klf4, and cMyc). (**E**) Proliferating progenitors (PPs) and newborn neurons (NNs) isolated from brains of mouse embryos. Horizontal lines delineate medians with mean absolute deviation (MAD) as error, datapoints represent medians of the individual replicates, the number of independent biological replicates is indicated below each box. Statistical analysis was performed using generalized linear mixed effects model. Data presented in (A) and (D) were previously published in ref (*30*) and (*31*), respectively.

The second system included small-cell and adenocarcinoma cell lines originating from human intestine, lung, and stomach. Consistently across tissues, small cell-carcinoma cells had a lower apparent Young's modulus compared to their adenocarcinoma counterparts (Fig. 2B). Small-cell carcinomas have comparatively small cell sizes, short doubling times and high metastatic potential, all connected with poor clinical prognosis in patients (*41*, *42*). Apart from the main transcriptomic dataset for the carcinoma project, in which all mechanically characterized cell lines are represented (FANTOM5 (*34*)), we collected three additional transcriptomic datasets generated with different expression profiling techniques (RNA-Seq or microarray profiling), and originating from different groups: Cancer Cell Line Encyclopedia (CCLE) microarray (*35*), CCLE RNA-Seq (*36*), and Genentech (*37*) (see Table 1 for overview). In the third studied system, a non-tumorigenic breast epithelium MCF10A cell line bearing single-allele oncogenic mutation H1047R in the catalytic subunit alpha of the phosphatidylinositol-4,5-bisphosphate 3-kinase (PIK3CA) (*38*, *43*) showed increased stiffness compared to the wild type control (Fig. 2C). H1047R mutation causes constitutive activation of PIK3CA and an aberrant triggering of the PI3K–AKT–mTOR signaling pathway leading to growth factor-independent proliferation (*44*, *45*). In the fourth system, the fuzzy-colony forming (F-class) state of iPSCs had a lower stiffness as compared to the bone-fide compact-colony forming (C-class) state (*31*) (Fig. 2D). C-class cells establish endogenous expression of reprogramming factors at moderate levels towards the end of reprogramming, while F-class cells depend on the ectopic expression of the pluripotency factors and are characterized by a fast proliferation rate (*39*). Finally, we characterized two stages of developing neurons isolated from embryonic mouse brain (*40*), and observed that the newborn neurons (NNs) had higher apparent Young's moduli than proliferating progenitors (PPs) (Fig. 2E). Cell areas and deformations used for Young's modulus extraction for all datasets are visualized in Fig. S2.

The mechano-transcriptomic datasets collected within the framework of our study (Table 1) represent a broad spectrum of biological systems encompassing distinct cell states associated with mechanical phenotype changes. The included systems come from two

different species (human and mouse), several tissues (brain, intestine, lung, stomach, breast, as well as embryonic tissue) and are associated with processes ranging from cancerogenic transformations to cell morphogenesis. This high diversity is important for focusing the analysis on genes universally connected to the change in mechanical properties, rather than on genes specific for processes captured by individual datasets.

Discriminative network analysis on prediction datasets

After characterizing the mechanical phenotype of the cell states, we set out to use the accompanying transcriptomic data to elucidate genes associated with the mechanical phenotype changes across the different model systems. To this end, we utilized a method for inferring phenotype-associated functional network modules from omics datasets termed PC-Corr (*29*). PC-Corr was performed individually on two prediction datasets, and the obtained results were overlayed to derive a conserved network module. Owing to the combination of Pearson's correlation and the discriminative information included in the PC loadings, the PC-corr analysis does not only consider gene co-expression — as is the case for classical co-expression network analysis — but also incorporates the relative relevance of each feature for discriminating between two or more conditions; in our case, the conditions representing soft and stiff phenotypes. The overlaying of the results from two different datasets allows for a multi-view analysis (utilizing multiple sets of features) and effectively merges the information from two different biological systems.

For the network construction, we chose two datasets that originate from different species, concern unrelated biological processes, and have high number of samples included in the transcriptional analysis: human glioblastoma and murine iPSCs (Table 1). PC-corr analysis was performed on these prediction datasets individually using a subset of transcripts at which the two datasets intersect (Fig. 3A). First, the 9,452 unique genes from the intersection were used to perform principal component analysis (PCA) (Fig. 3, B and C). Next, the PC loadings for the component showing good separation between the different cell states (PC1 for both of presented datasets) were normalized and scaled (see Methods for details). The

processed PC loadings, V, were then combined with Pearson's correlation coefficients, c, to obtain a $PC\text{-}corr$ value for each pair of genes i, j for every n-th dataset according to the following formula:

$$PC\text{-}corr_{i,j}^{n} = \mathrm{sgn}\left(c_{i,j}^{n}\right) \min\left(|V_i^n|, |V_j^n|, |c_{i,j}^{n}|\right). \quad (1)$$

The sign of the $PC\text{-}corr$ value corresponds to the correlated (positive) or anti-correlated (negative) expression of genes i, j, and the magnitude of $PC\text{-}corr$ conveys the combined information about the strength of the expression correlation and the contribution of the individual genes to the phenotype-based separation of samples along the principal component.

To merge the PC-corr results obtained for the individual prediction datasets, a combined $PC\text{-}corr$ value, $PC\text{-}corr_{i,j}^{comb}$, was calculated either as a mean or as a minimum of the individual $PC\text{-}corr$ values. For n datasets:

$$PC\text{-}corr_{i,j}^{comb} = \begin{cases} \delta_{i,j} \ \dfrac{1}{N} \displaystyle\sum_{n=1}^{N} |PC\text{-}corr_{i,j}^{n}| \\ \delta_{i,j} \min\left(|PC\text{-}corr_{i,j}^{1}|, \dots, |PC\text{-}corr_{i,j}^{n}| \right) \end{cases} \quad (2)$$

where $\delta_{i,j} \in \{-1,1\}$ defines the sign of $PC\text{-}corr_{i,j}^{comb}$, and is equal to the mode of $PC\text{-}corr_{i,j}$ signs over all individual datasets. In our implementation on two datasets, gene pairs with opposing $PC\text{-}corr$ signs were masked by setting their $PC\text{-}corr^{comb}$ values to zero.

To obtain the network of putative target genes, a cut-off was applied to the absolute $PC\text{-}corr^{comb}$ value. We explored several cut-off strategies in order to obtain a wide overview of the meaningful conserved network structures. By looking at $PC\text{-}corr^{comb}$ calculated as mean and setting the threshold for its absolute value to 0.75, we obtained a network of 29 nodes connected by 30 edges (Fig. 3E). The edges describe the connection between the genes in the network and their thickness is defined by the $PC\text{-}corr^{comb}$ values (Table S1). The node colors reflect the strength of the contribution of individual genes to the separation of the different classes as described by the mean of the processed PC loadings V.

The obtained network can be made more conservative by using the minimum $PC\text{-}corr^{comb}$ instead of the mean, or by changing the cut-off value. Utilizing the $PC\text{-}corr^{comb}$ calculated as minimum value (Table S2) and setting the cut-off value to 0.70, we obtained a network with 22 nodes connected by 29 edges (Fig. 3F). Increasing the cut-off value to 0.75 resulted in a network of 9 genes connected by 12 edges (Fig. 3G). The list of genes from the three networks presented in Fig. 3, E to G, together with their full names and processed PC loading values, is presented in Table S3.

We performed gene ontology enrichment analysis for biological processes on the nodes of the network presented in Fig. 3G, as well as the union of all nodes presented in Fig. 3, E to G (Fig. S3). The top two significantly enriched terms in the 9-gene set were the negative regulation of transcription by polymerase II (GO:000122) and negative regulation of endothelial cell proliferation (GO:0001937). In the 34-gene set, apart from a broad term of signal transduction (GO:0007165), the significantly enriched terms included negative regulation of transcription by polymerase II (GO:000122), regulation of cell growth (GO:0001558), and negative regulation of cell proliferation (GO:0008285), among others. The fact that these GO terms are not obviously related to cell mechanics might be an indicator that the association of the identified genes with cell mechanics is relative unknown, and that our mechanomics approach can identify such associations de novo. The aforementioned categories included mostly genes showing high expression in the stiff states. Since these highly expressed genes are associated with negative regulation of growth and transcription, our results point towards reduced transcriptional activity and reduced growth/proliferation in stiff compared to soft cells.

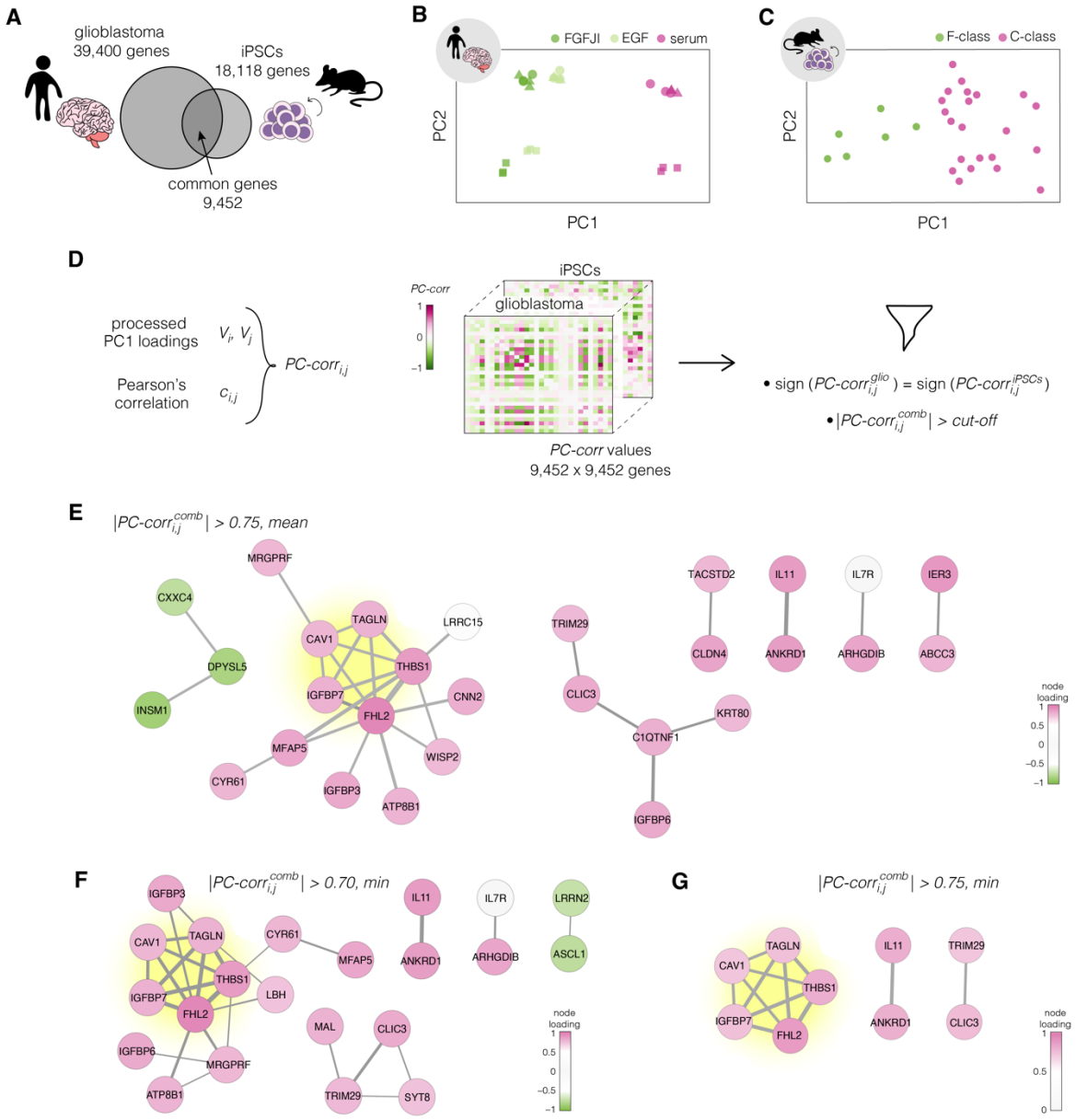

Fig. 3. Identification of putative targets involved in cell mechanics regulation. (**A**) Glioblastoma and iPSC transcriptomes used for the target prediction intersect at 9,452 genes. (**B**, **C**) PCA separation along two first principal components of the mechanically distinct cell states in the glioblastoma (B) and iPSC (C) datasets. The analysis was performed using the gene expression data from the intersection presented in (A). (**D**) Schematic representation of PC-corr analysis and the combination of the PC-corr results for two systems. (**E** to **G**) Gene networks based on filtering gene pairs by the combined *PC-corr* score. The presented networks were obtained by setting the cut-off value to 0.75, when using the mean PC-corr approach (E), and to 0.70 (F) and 0.75 (G), when using the minimum value approach. In (E to G) edge thicknesses represent the $|PC\text{-}corr^{comb}|$ (Table S1 and S2) and the colors of the nodes represent the average processed PC loadings (Table S3).

The identified conserved functional network module comprises five genes

Regardless of the strategy chosen for the selection of the network-building gene pairs, a strongly interconnected module of 5 genes (Table 2) — highlighted in yellow in Fig. 3, E to G — emerged. We focused on the five genes from this conserved network module as putative targets for regulating cell mechanics: *CAV1*, *FHL2*, *IGFBP7*, *TAGLN*, and *THBS1*.

Table 2. List of identified target genes comprising the conserved module.

symbol	gene description	HGNC ID	MGI ID
CAV1	caveolin 1	HGNC:1527	MGI:102709
FHL2	four and a half LIM domains 2	HGNC:3703	MGI:1338762
IGFBP7	insulin like growth factor binding protein 7	HGNC:5476	MGI:1352480
TAGLN	transgelin	HGNC:11553	MGI:106012
THBS1	thrombospondin 1	HGNC:11785	MGI:98737

Caveolin-1, CAV1, is a protein most prominently known for its role as a structural component of caveolae. Caveolae are small cup-shaped invaginations in the cell membrane that are involved, among other functions, in the mechanoprotective mechanism of buffering the plasma membrane tension (*46*, *47*). Recent data suggests that CAV1 can also confer its mechanoprotective role independently of caveolae (*48*). Apart from membrane organization and membrane domain scaffolding, CAV1 plays a role in an array of regulatory functions such as metabolic regulation or Rho-signalling (*47*). The second identified target, four and a half LIM domains 2, FHL2, is a multifaceted LIM domain protein with many binding partners and a transcription factor activity (*49*). FHL2 has recently been shown to remain bound to actin filaments under high tension, and be shuttled to the nucleus under low cytoskeletal tension (*50*, *51*) — a property conserved among many LIM domain-containing proteins (*51*, *52*). The third target, Insulin-like growth factor binding protein 7, IGFBP7, is a secreted protein implicated in a variety of cancers. It is involved in the regulation of processes such as cell

proliferation, adhesion, and senescence (*53*). Transgelin, TGLN, is an actin-binding protein whose expression is up-regulated by high cytoskeletal tension (*54*) and is also known to play a role in cancer (*55*). Finally, thrombospondin 1, THBS1, is a matricellular, calcium-binding glycoprotein that mediates cell-cell and cell-matrix adhesions and has many regulatory functions (*56*, *57*).

Before validating the performance of the five target genes, we inspected their expression across the divergent cell states in the collected datasets. The target genes show clear differences in expression levels between the soft and stiff cell states and provide for clustering of the samples corresponding to different cell stiffnesses in both prediction and validation datasets (Fig. 4). The relation between normalized apparent Young's modulus change and fold-change in the expression of the target genes is presented in Fig. S4. Of note, the direction of changes in the expression levels between the soft and stiff cell states in the validation datasets was not always following the same direction (Fig. 4, C to F, Fig. S4). This suggests that the genes associated with cell mechanics may not have a monotonic relationship with cell stiffness, but rather are characterized by different expression regimes in which the expression change in opposite directions can have the same effect on cell stiffness.

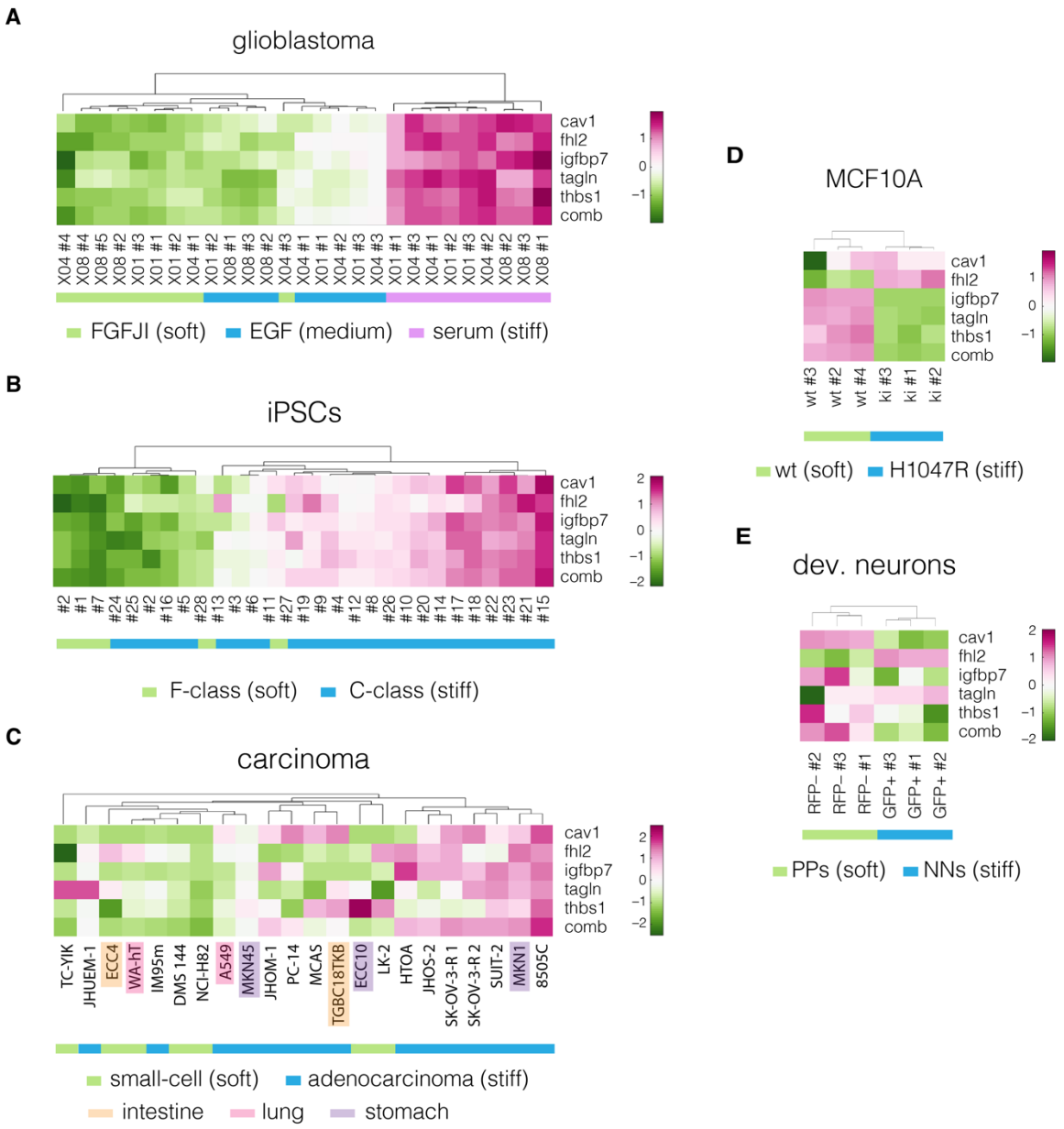

Fig. 4. Expression of identified target genes in the prediction and validation datasets. Panels show unsupervised clustering heat maps of expression data from transcriptomic datasets corresponding to the following systems: (**A**) glioblastoma, (**B**) induced pluripotent stem cells (iPSCs), (**C**) carcinoma (FANTOM5), (**D**) non-tumorigenic breast epithelia MCF10A, and (**E**) developing neurons (dev. neurons). Comb – combinatorial marker, wt – wild type, PPs – proliferating progenitors, NNs – newborn neurons. Clustering was performed using *clustergram* function in *MATLAB* (*R2020a,* MathWorks) on log-normalized expression data.

Universality, specificity, and trustworthiness of the identified markers

Next, we validated whether the five identified genes individually, as well as their association into a unique combinatorial marker (computed as the mean of the five log-normalized genes, see Methods), are universal and specific markers of cell mechanics. To assess that, we tested three hypotheses using combinations of transcriptomic data from six validation datasets as detailed in Table 3. The classification performance of each marker was assessed using the area under the curve of the receiver-operator characteristics (AUC-ROC) (*58*), which takes values from 0 to 1, with 1 corresponding to a perfect classifier and 0.5 to a random classifier. Importantly, for each hypothesis multiple datasets were used, and the discriminative performance was assessed in a joint multiview way by looking at the minimum value of AUC-ROC across multiple comparisons.

Table 3. Overview of the hypotheses and datasets used for validating universality and specificity of obtained markers. Hypotheses are listed in the column headings. Under every hypothesis, sample groups used for the hypothesis testing are listed. Numbers of samples used in every group are indicated in brackets.

positive hypothesis I markers are discriminative of samples with stiff/soft mechanical phenotype independent of the studied biological system	**positive hypothesis II** markers are discriminative of samples with stiff/soft independent of data source	**negative hypothesis** markers are discriminative of samples from different tissue of origin (but with no mechanical difference)
Carcinoma | FANTOM5 1. small-cell (6) vs adeno (6) (lung, intestine and stomach) **MCF10A** 2. wt (3) vs H1047R (3) **developing neurons** 3. PPs (3) vs NNs (3)	**Carcinoma | CCLE microarray** 1. small-cell (51) vs adeno (49) (lung) **Carcinoma | CCLE RNA-Seq** 2. small-cell (51) vs adeno (77) (lung) **Carcinoma | Genentech (RNA-Seq)** 3. small-cell (30) vs adeno (38) (lung)	**Carcinoma | CCLE microarray** 1. lung (49) vs stomach (19) (adeno) 2. large intestine (43) vs stomach (19) (adeno) **Carcinoma | Genentech (RNA-Seq)** 3. lung (38) vs stomach (14) (adeno)

We first tested whether the obtained markers are universal across systems of different biological origin (positive hypothesis I) by estimating their ability to discriminate between stiff and soft cell phenotypes in three validation datasets: developing neurons (mouse), carcinoma cell lines originating from three tissues (human), and MCF10A (human) (Table 4). Particularly high minimum AUC-ROC values (≥ 0.78) were obtained for CAV1, FHL2, and TAGLN, and the combinatorial marker outperformed all of the individual genes with a minimum AUC-ROC of 0.97. The ROC curves for individual datasets are presented in Fig. S5.

Table 4. Validation of identified target genes and the combinatorial marker. Minimum AUC-ROC (min AUC-ROC) and JVT *p* values are reporter for the two positive hypotheses and one negative hypothesis for each target genes and the combinatorial marker (comb). The specific datasets and comparisons used for testing of each hypothesis are specified in Table 3 above. The results presented in this table can be reproducible using the code and data available under the GitHub link reported in the methods section.

	measure	*CAV1*	*FHL2*	*IGFBP7*	*TAGLN*	*THBS1*	comb
positive hypothesis I	min AUC-ROC	0.78	0.89	0.67	0.78	0.56	**0.97**
	JVT *p* value	0.14	0.04	0.30	0.14	0.81	**0.01**
positive hypothesis II	min AUC-ROC	0.89	0.88	0.73	0.56	0.86	**0.92**
	JVT *p* value	0.02	0.03	0.19	0.59	0.04	**0.01**
negative hypothesis	min AUC-ROC	0.54	0.51	0.51	0.52	0.61	**0.51**
	JVT *p* value	0.40	0.76	0.90	0.61	0.06	**0.91**

Next, we tested whether the identified markers provide good sample classification across similar datasets obtained from different sources (positive hypothesis II). For this purpose, we used three carcinoma datasets that were generated by two different research group using either microarray or RNA-Seq (see Table 1 and 3). Within these datasets, we looked at the discrimination between the small-cell and adenocarcinoma samples from lung. This choice was dictated by the highest number of available samples from this tissue across the datasets.

Also here, the multiview AUC-ROC values were high, reaching 0.89 for CAV1, 0.88 for FHL2, and 0.86 for THBS1. The combinatorial marker had an AUC-ROC value of 0.92.

To assess whether the predicted markers are specific to the mechanical phenotype, we tested their performance in classification of the adenocarcinoma samples grouped by the tissue they were derived from (negative hypothesis). This groups did not show clear mechanical differences (Fig. 2B) For the combinatorial marker, the min AUC-ROC value was equivalent to a random classifier (0.51), and for the individual markers reached values between 0.51 and 0.65 (Table 4). Since the discriminative power of the obtained markers vanished (reached AUC-ROC close to 0.50 corresponding to a random classifier) when tested on groups that do not encompass cell mechanic phenotype difference, we can conclude that the identified markers are specific to the mechanical phenotype.

Finally, to test the trustworthiness of obtained markers, we evaluated how easy it is to generate markers with equivalent discriminative power at random. For that purpose, we devised a novel methodology called joint-view trustworthiness (JVT). JVT is a resampling technique that creates a null model distribution according to which an empirical *p* value is computed to evaluate the probability to sample at random a marker that offers a joint multview discrimination equal or better to the one of the predicted markers (see Methods for details). A low JVT *p* value (< 0.05 significant level) means that it is rare to randomly generate a joint multiview marker with performance equal or better than the tested one. As summarized in Table 4, the combinatorial marker had remarkably low JVT *p* values ($p = 0.01$) in positive hypotheses I and II, i.e., it is very unlikely to generate a similarly performing combinatorial marker at random. Conversely, in the negative hypothesis, the JVT *p* value of the combinatorial marker is not significant ($p = 0.91$). The performance of the tested genes individually was varied, with FHL2 showing a significant JVT *p* value in positive hypothesis I, and FHL2, CAV1 and THBS1 reaching significant JVT *p* values in positive hypothesis II. It is important to note that our implementation of JVT is conservative, as we consider the minimum discriminative performance on multiple datasets. This may lead to

underestimating the performance of individual markers. In sum, the results provided in Table 3 pointed towards CAV1 and FHL2 as promising markers of the mechanical phenotype.

Perturbing expression levels of CAV1 changes cells stiffness

We decided to focus our attention on CAV1 as a potential target for modulating mechanical properties of cells, as it has previously been linked to processes intertwined with cell mechanics. In the context of mechanosensing, CAV1 is known to facilitate buffering of the membrane tension (*46*), play a role in β1-inegrin-dependent mechanotransduction (*59*) and modulate the mechanotransduction in response to substrate stiffness (*60*). CAV1 is also intimately linked with actin cytoskeleton — it was shown to be involved in cross-talk with Rho-signaling and actin cytoskeleton regulation (*47*, *61–63*), filamin A-mediated interactions with actin filaments (*64*), and co-localization with peripheral actin (*65*). Yet, the data supporting the direct impact of CAV1 on mechanical properties of cells is, so far, limited (*63*, *66–68*).

In most of the mechano-transcriptomic datasets considered in our study, the increase in apparent Young's modulus was accompanied by an increase in CAV1 levels (Fig. S4A), corroborating with previous reports (*66*, *69*). Additionally, we observed that mouse embryonic fibroblasts isolated from CAV1 knock out mice (CAV1KO) are softer than the wild type cells (WT) (Fig. S6). Thus, we set out to test weather artificially decreasing the levels of CAV1 results in cell softening, and conversely, increasing the level of CAV1 in higher cell stiffness. To this end, we perturbed the levels of CAV1 in the cell lines representing two intestine carcinoma types: ECC4, the small-cell carcinoma with a comparably soft phenotype, and TGBC18TKB (TGBC), the adenocarcinoma with a comparatively stiff phenotype. We confirmed that TGBC cells have a higher level of CAV1 as compared to ECC4 on a protein level (Fig. 5A) and that they are characterized by a stiffer phenotype, not only when measured with RT-DC (Fig. 2B, Fig. 5B), but also with atomic force microscopy (AFM) using both standard indentation experiments (Fig. 5C), as well as oscillatory measurements at different frequencies, referred to as AFM microrheology (Fig. 5D).

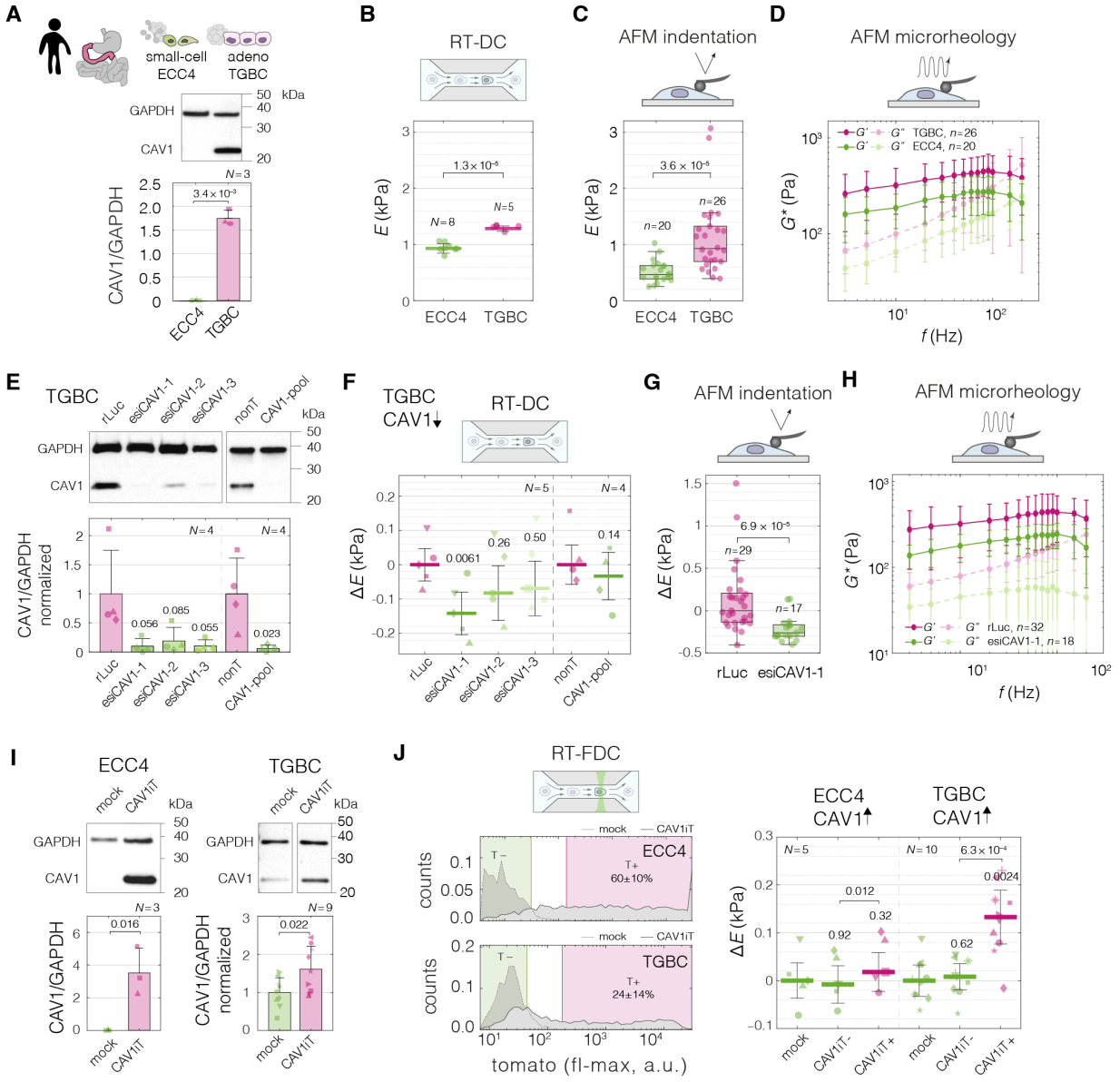

Fig. 5. Perturbing levels of CAV1 affects the mechanical phenotype of intestine carcinoma cells. (**A**) CAV1 levels in small-cell (ECC4) and adenocarcinoma (TGBC) cell lines from intestine. (**B** to **D**) Mechanical phenotype of ECC4 and TGBC cells measured with RT-DC (B**,** as in Fig. 2B), AFM indentation (C), and AFM microrheology (D). (**E**) Verification of CAV1 knock-down in TGBC cells using two knock-down system: three esiRNA constructs (esiCAV1-1. esiCAV1-1, and esiCAV1-3 with rLuc as a control), and pooled siRNA mixture (CAV1-pool with non-targeting mixture (nonT) as a control). (**F** to **H**) Mechanical phenotype change of TGBC cells upon CAV1 knock-down as measured by RT-DC (F), AFM indentation (G), and AFM microrheology (H). (**I**) Verification of transient CAV1 overexpression in ECC4 and TGBC cells. (**J**) Mechanical phenotype change of ECC4 and TGBC cells upon CAV1 overexpression as measured by RT-FDC. Gating for fluorescence positive and negative cells based on

dTomato expression in ECC4 (top) and TGBC (bottom) cells (left-hand side). Fluorescence positive cells correspond to cells expressing CAV1-IRES-dTomato (CAV1iT). For comparison, mock transfection sample is shown in the background (mock). Apparent Young's modulus changes of ECC4 and TGBC cells upon CAV1 overexpression (right-hand side). CAV1iT- and CAV1T+ are dTomato negative and positive cells, respectively. For protein quantification in (A), (E), and (I), representative Western blots (top) as well as quantification of specified replicate numbers *N* (bottom) are shown. In (B), (F) and (J), horizontal lines delineate medians with mean absolute deviation (MAD) as error, datapoints represent medians of *N* experiment replicates, statistical analysis was performed using generalized linear mixed effects model. In (C) and (G), box plots spread from 25th to 75th percentiles with a line at the median, whiskers span 1.5 × interquartile range (IQR), individual datapoints correspond to values obtained for *n* individual cells, statistical analysis was performed using two sample two-sided Wilcoxon rank sum test. In (D) and (H), datapoints correspond to means ± standard deviation of all measurements at given oscillation frequencies for *n* cells. Lines connecting datapoints serve as guides for the eye. E – apparent Young's modulus, G^* – complex shear modulus, ΔE – apparent Young's modulus change relative to respective control measurements.

To decrease the levels of CAV1 in the TGBC cells, we performed knock-down experiments using two RNA interference (RNAi) systems, endoribonuclease-prepared siRNA (esiRNA) targeting three different parts of CAV1 transcript (esiCAV1-1, esiCAV1-2, and esiCAV1-3), and a pool of conventional siRNAs (CAV1-pool) (Fig. 5E). All the RNAi approaches resulted in the decrease of the apparent Young's modulus of TGBC cells as measured by RT-DC (Fig. 5F, Fig. S7, C and D). The most prominent effect was observed using esiCAV1-1. We further confirmed that CAV1 knock-down with esiCAV1-1 resulted in decreased stiffness of TGBC cells using AFM indentation (Fig. 5F) and microrheology (Fig. 5G).

To investigate the influence of increased CAV1 levels on cell stiffness, we performed transient overexpression experiments of CAV1 with a dTomato reporter under independent ribosomal entry site, IRES, (CAV1iT) in both ECC4 and TGBC cell lines. At 72 hours post transfection, we observed elevated levels of CAV1 in both cell lines on a protein level in bulk (Fig. 5I). Since in the transient overexpression experiments not all of the cells are transfected, we leveraged the possibility to monitor the fluorescence of single cells in parallel with their mechanical phenotype offered by real-time fluorescence and deformability cytometry (RT-FDC) (*32*) to gate for the fluorescence-positive cells (T+, gate marked in magenta in Fig. 5J). The fluorescence-positive cells in the CAV1-transfected sample, CAV1iT+, showed higher

apparent Young's moduli as compared to fluorescence-negative cells in both control sample (mock) and CAV1-transfected sample (CAV1iT–, internal control) (Fig. 5J, Fig. S7, C and D). The effect was observed in ECC4 as well as TGBC cells. However, it was more pronounced in the TGBC cells, suggesting that the cells may be more responsive to the artificial increase in CAV1 levels when natively expressing a basal level of this protein.

Finally, we performed CAV1 perturbation experiments in a breast epithelial cell model of cancerous transformation, MCF10A-ER-Src cells, in which the Src proto-oncogene can be induced by treatment with tamoxifen (TAM). As previously shown, TAM addition triggers Src phosphorylation and cellular transformation (*70*), which is associated with F-actin cytoskeletal changes and, after a transient stiffening, the acquisition of a soft phenotype evident at 36 hours post induction (*71*). We inspected a previously published microarray dataset and determined that the expression of CAV1 diminishes over time after TAM treatment (*72*) (Fig. 6A). We then showed that the decrease of CAV1 could be observed at the protein level 72 hours post induction (Fig. 6B), a timepoint at which the TAM-induced MCF10A-ER-Src cells show a significant decrease in cell stiffness (ref (*71*) and Fig. 6C). We next showed that decreasing the level of CAV1 by knock-down caused a decrease in stiffness of uninduced MCF10A-ER-Src cells similar to that caused by TAM induction (Fig. 6D). Finally, we performed an inverse experiment, in which we rescued the CAV1 levels in TAM-induced MCF10A-ER-Src cells by transient overexpression. The cells with CAV1 overexpression showed a stiff phenotype, corresponding to the one of uninduced cells (Fig. 6E).

Taken together, the results obtained with the intestine carcinoma cell lines and MCF10A-ER-Src cells show that CAV1 not only correlates with, but also is causative of mechanical phenotype change.

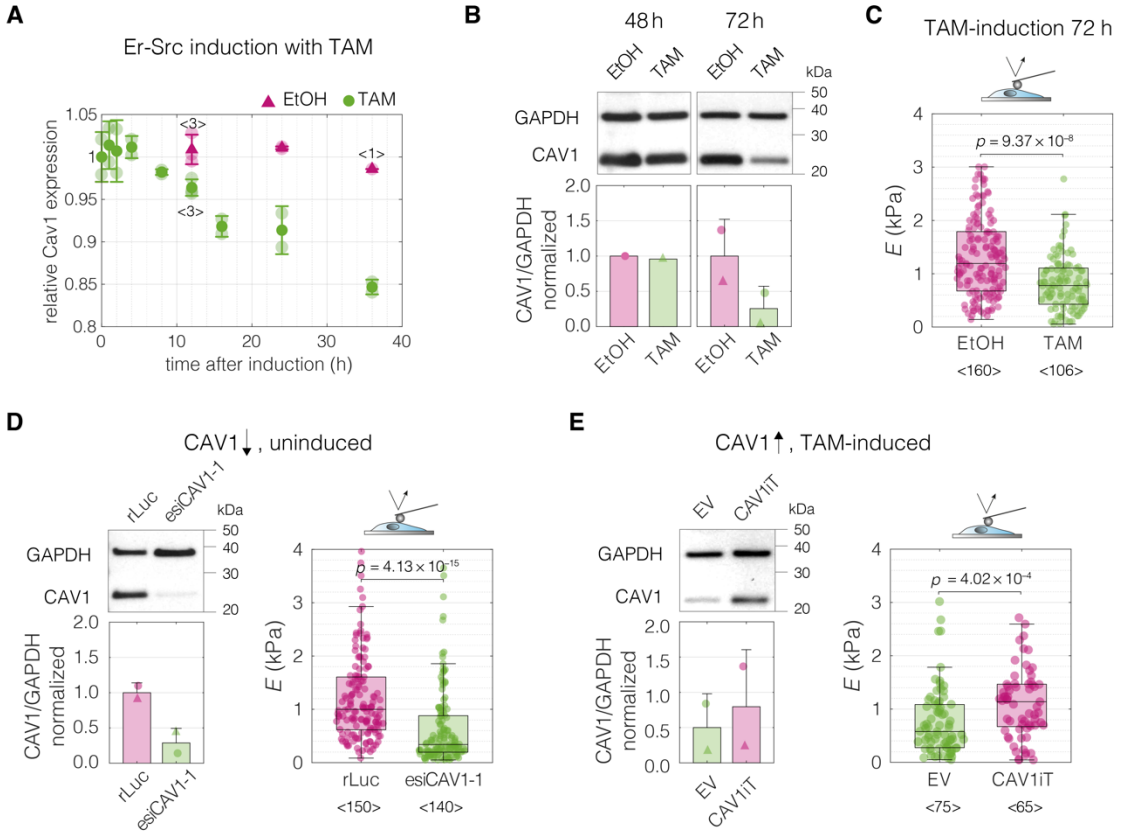

Fig. 6. Perturbations of CAV1 levels in MCF10A-ER-Src cells result in cell stiffness changes. (**A**) Inducing transformation of MCF10A- ER-Src cells by tamoxifen (TAM) treatment, as opposed to vehicle control (ethanol, EtOH), causes a decrease of CAV1 expression over time, as captured by microarray analysis (GEO accession number: GSE17941, data previously published in ()). Datapoints with error bars represent means ± 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 to E) spread from 25th to 75th percentiles with a line at the median, whiskers span 1.5 × 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.

Discussion

The mechanical phenotype of cells is recognized as a hallmark of many physiological and pathological processes. Understanding how to control it is a necessary next step that will facilitate exploring the impact of cell mechanics perturbations on cell and tissue function (*3*). The increasing availability of transcriptional profiles accompanying cell state changes has recently been complemented by the ease of screening for mechanical phenotypes of cells thanks to the advent of high-throughput microfluidic methods (*73*). This provides an opportunity for data-driven identification of genes associated with the mechanical cell phenotype change in a hypothesis-free manner. Here we leveraged this opportunity by performing discriminative network analysis on transcriptomes associated with mechanical phenotype changes to elucidate a conserved module of five genes potentially involved in cell mechanical phenotype regulation. We provided strong evidence that the inferred conserved functional network module contains an ensemble of five genes that, in particular when combined in a unique combinatorial marker, are universal, specific and trustworthy markers of mechanical phenotype. We further demonstrated on the example of a selected marker gene, CAV1, that its experimental up- and downregulation impacts the stiffness of the measured cells. This demonstrates that the level of CAV1 not only correlates with, but also is causative of mechanical phenotype change.

The workflow presented here is a blueprint for data-driven discovery of cell mechanics markers that can serves as targets for modulating cell mechanical properties. Its key features are the hypothesis-free modus operandi and the integration of information from different biological systems, that allows to focus on genes that play a relatively general role in cell mechanics rather than on genes specific to the individual experimental models. Noteworthy, by including the PC loadings in the scores used for thresholding, the PC-corr method implemented for network analysis in our study offers a multivariate alternative to classical co-expression analysis, that highlights not only the correlation between the genes but also their relative importance for separating samples based on their mechanical phenotype.

Despite its simplicity, PC-corr offers a robust performance on different types of omics data, and has already proven its efficacy in several studies (*29*, *30*, *74*).

Among the target genes elucidated in our analysis, we did not observe enrichment of gene ontology terms related to actin cytoskeleton organization, actomyosin contractility, or cell migration — processes that are typically associated with cell mechanics (Fig. S3). This can be partially explained by looking at the mRNA rather than the protein level, its supramolecular assembly, activation state or localization. Upon closer inspection of the obtained gene targets, we found some links connecting them with cell mechanics in the literature. As indicated above, CAV1 has been shown to be involved in cross-talk with Rho-signaling and actin-related processes, as well as physical interactions with actin (*47*, *61–65*). It is thus conceivable that CAV1 is involved in cell mechanics regulation via its influence on the actin cytoskeleton and its contractility. Furthermore, CAV1 is known to modulate the activation of transcriptional cofactor yes-associated protein, YAP, in response to changes in stiffness of cell substrate (*60*) and in the mechanical stretch-induced mesothelial to mesenchymal transition (*75*). YAP is an established transducer of not only various mechanical stimuli, but also of cell shape and the changes in the actin cytoskeleton tension (*76*), the latter being an important determinant of cell stiffness. Conversely, YAP is an essential co-activator of CAV1 expression (*77*). In the extended networks (Fig. 3, E and F), we found three further genes that are identified (CYR61, ANKRD1) (*78*, *79*) or implicated (THBS1) (*76*) as transcriptional targets of YAP. The next identified marker gene, transgelin, TGLN (also known as SM22α) is an actin-binding protein, that stabilizes actin filaments and is positively correlated with cytoskeletal tension (*80*). Transgelin is a member of the calponin protein family, one further member of which, calponin 2, CNN2, is present in the broader sets of genes identified in this study (Fig. 3, E and F, Table S3). The expression of calponin 2, likewise, stabilizes actin filaments and is increased in cells with high cytoskeletal tension (*81*). Finally, FHL2 is a transcriptional coactivator that is found, together with other LIM domain protein families such as zyxin and paxillin, to localize to actin filaments that are under stress (*50–52*). When the cytoskeletal

tension is low, FHL2 translocates to the nucleus, thus serving as a nuclear transducer of actomyosin contractility (*50*).

To our knowledge, there are no prior studies that aim at identifying gene signatures associated with mechanical phenotype changes, in particular across different cell types. There are, however, several studies that investigated changes in expression upon exposure of specific cell types to mechanical stimuli such as compression (*82*, *83*) or mechanical stretch (*23*, *75*, *84*), and one study that investigated difference in expression profiles between stiffer and softer cells sorted from the same population (*85*). Even though the studies concerned with response to mechanical stimuli answer a fundamentally different question (how gene expression changes upon exposure to external forces vs which genes are expressed in cells of different mechanical phenotype), we did observe some similarities in the identified genes. For example, in the differentially expressed genes identified in the lung epithelia exposed to compression (*82*), three genes from our module overlapped with the immediate response (CAV1, FHL2, TGLN) and four with the long-term one (CAV1, FHL2, TGLN, THBS1). We speculate that this substantial overlap is caused by the cells undergoing change in their stiffness during the response to compression (and concomitant unjamming transition).

As seen from the example of the target genes included in the conserved module, their change is correlated with cell mechanics across all datasets, but it does not always follow the same trend (Figs. 4 and S4). This non-monotonic relationship between gene expression and the mechanical phenotype change suggests that there may be different regimes at which the expression change in the same direction has an opposite effect on the property of interest. Furthermore, the effect of expression change may be contextual and depend on the state of cells. This observation carries some parallels to the role of several of our target genes in cancer progression. For example, CAV1 has been indicated as both promoting and suppressing cancer progression in a variety of tissues. One way in which this can be reconciled is that the change in CAV1 expression may have different roles depending on the stage of caner progression (*61*, *86*, *87*). A similar ambiguity of their role in cancer progression was indicated for THBS1 (*57*) and IGFBP7 (*53*). Of note, a non-monotonic cell stiffness

response has also been described for treatments with actin-disrupting drugs. For example, treating cells with Latrunculin B makes cells progressively more deformable up to a certain concentration, beyond which the cells become less deformable again and eventually even stiffer than non-treated cells (see (*73*) and discussion therein for more references). Apart from characterizing the response regimes, it will be also important to consider the temporal dynamics of cell response to the change in expression of a given gene. Trying to push the cell out of its equilibrium may cause the system to respond actively to counterbalance the induced change, which, in turn, may lead to oscillations in both expression levels of manipulated protein and its effectors, as well as the mechanical properties of the cell.

Among all different types of omics data, looking at the transcriptome is advantageous and disadvantageous at the same time. Its limitation is that mRNA levels do not necessarily reflect protein content in cells. Furthermore, for many proteins it is not the absolute level that has a functional relevance, but rather the protein activation by, for example, phosphorylation or binding with co-activators, or its localization. However, identifying the players at the transcriptome level has the advantage of easy implementation in perturbation experiments with established genetic tools, such as CRISPR-Cas9 technology or RNAi. Furthermore, our analysis framework is readily applicable to other types of omics data, including proteomic, metabolomic, lipidomic, or glycomic data, the analysis of which would complement our study and provide different insights into the regulation of cell mechanics. Lipidomic data, for example, could reveal possible contributors to cell mechanics related to the composition of the cell membrane.

For the approaches such as the one pioneered in this study to flourish, it is necessary that the mechanical datasets become routinely published and annotated in a manner similar to omics datasets. With the recent advent of high-throughput cell mechanical characterization techniques, such as deformability cytometry methods (*73*), the establishment of a database for cell mechanics gains immediate relevance. In our group alone, within the timespan of five years since the RT-DC method was originally published (*32*), we have accumulated over 100,000 individual mechanical characterization experiments, comprising roughly a billion of

single cells measured. Once a vast number of mechanics datasets connected to omics profiles is available, it will be straightforward to develop a next generation artificial intelligence algorithm predicting cell stiffness from given omics profiles. Apart from analyzing divergent cell states, the search for mechanical regulators could be complemented by looking into omics data of cells from unimodal populations sorted by their mechanical properties — a pursuit that with the advent of high-throughput methods for mechanics-based sorting of cells, such as sorting RT-DC (*88*, *89*) or passive filtration-based approaches (*85*), becomes a realistic objective.

In conclusion, this work brings together machine learning-based discriminative network analysis and high-throughput mechanical phenotyping to establish a blueprint workflow for data-driven de novo identification of genes involved in the regulation of cell mechanics. Ultimately, identifying ways to tune the mechanical properties on demand will enable turning cell mechanics from a correlative phenomenological parameter to a controllable property. Such control will, in turn, allow us to interfere with important processes such as tissue morphogenesis, cell migration, or circulation through vasculature.

Methods

CELL CULTURE

Glioblastoma cell lines

The glioblastoma dataset contained three primary human brain tumor cell lines (X01, X04, and X08) in three distinct signaling states. The cells were cultured and characterized within a framework of a previous study (*30*). In brief, the three signaling states characterized by low, medium, and high activation of STAT3-Ser/Hes3 signaling axis, were maintained by growth media containing fetal bovine serum (serum), epidermal growth factor (EGF), or basic fibroblast growth factor combined with a JAK inhibitor (FGFJI), respectively. Upon thawing, cells were expanded in a serum-free DMEM/F12 medium (10-090-CV, Mediatech, Corning, NY, USA) containing N2 supplement and 20 ng ml^{-1} EGF (R&D Systems, MN, USA) at 37°C in a 5% oxygen incubator. Each cell line was then plated into three separate flasks and cultured in the DMEM/F12 medium containing N2 supplement and additional supplementation of either serum (10%), EGF (20 ng ml^{-1}), or FGFJI (20 ng ml^{-1}, bFGF, R&D Systems; and 200 nM JAK inhibitor, Calbiochem, Merck Millipore, Germany). Cells were collected for mechanical characterization and RNA-Seq after 5-day exposure to the respective culture conditions (*30*).

Carcinoma cell lines

Small-cell and adenocarcinoma cell lines from intestine, stomach and lung were acquired from RIKEN BioResource Research Center, Japan (see Table S4 for the list of cell lines and media). Cells were cultured in growth media supplemented with 5% (TGBC) or 10% (rest) heat-inactivated fetal bovine serum (10270106, Gibco, ThermoFisher Scientific, MA, USA) and 100 U ml^{-1}/100 µg ml^{-1} penicillin/streptavidin (15140122, Gibco), at 37°C and 5% CO_2. Sub-culturing was performed using trypsin (25200072, Gibco). Cells were collected for mechanical characterization at 70% confluency. The RNA-Seq data was retrieved from FANTOM5 consortium (*34*). Additional transcriptomic datasets were retrieved from the CCLE project

(microarray (*35*) and RNA-Seq (*36*)) and from the study conducted by Genentech (*37*) (see Table 1 for overview).

MCF10A PIK3CA cell lines

MCF10A cell line with single-allele PIK3CA H1024R mutation was previously generated by homologous recombination by Horizon Discovery LTD, UK (*43*) and was kindly provided, together with an isogenic wild type (wt) control, by L.R. Stephens (Babraham Institute, UK). Cells used for mechanical characterization were cultured in DMEM/F12 medium (31330038, Gibco) supplemented with 5% horse serum (PAA Laboratories), 10 μg ml^{-1} insulin (I9278, Sigma-Aldrich, MO, USA), 0.2 μg ml^{-1} hydrocortisone (H0888, Sigma-Aldrich), 0.1 μg ml^{-1} cholera toxin (C8052, Sigma-Aldrich), and 100 U ml^{-1}/100 μg ml^{-1} penicillin/streptomycin (15140122, Gibco). The wt cells were additionally supplemented with 10 ng ml^{-1} EGF (E9644, Sigma-Aldrich), while mutant cell liens were maintained without EGF. Sub-confluent cells were collected for mechanical characterization using trypsin (25200056, Gibco). Mechanical data were collected from two biological replicates with three technical repetitions each. The RNA-Seq data was retrieved from a previous study (*38*), in which cells were cultured in a reduced medium (DMEM/F12 supplemented with 1% charcoal dextran treated fetal bovine serum, 0.2 μg ml^{-1} hydrocortisone and 0.1 μg ml^{-1} cholera toxin).

Induced pluripotent stem cells

F- and C-class iPSCs were derived through reprogramming of murine fetal neural progenitor cells with Tet-On system for doxycycline-inducible expression of OSKM (Oct4, Sox2, Klf4, cMyc) factors in a previous study (*31*). Both iPSCs classes were cultured on 0.1% gelatin-coated dishes in FCS/LIF medium (DMEM+Glutamax (61965059, Gibco), 15% fetal calf serum (Pansera ES, PAN-Biotech, Germany), 100 μM β-mercaptoethanol (PAN-Biotech), 2 mM L-glutamine, 1 mM sodium pyruvate, 1× nonessential amino acids, 15 ng ml^{-1} recombinant LIF (MPI-CBG, Dresden, Germany)). The F-class iPSCs were additionally supplemented with 1 μg ml^{-1} doxycycline, and the C-class iPSCs with a mixture of two inhibitors (2i): 1 μM MEK inhibitor (PD0325901, Calbiochem) and 3 μM GSK3 inhibitor (CH99021, Calbiochem). Cells

were passaged and harvested using 0.1% trypsin solution. The mechanical characterization was performed not earlier than at the 27th day of reprogramming (*31*). The microarray expression profiles were retrieved from a previous study, in which the F- and C-class iPSCs were derived from embryonic fibroblasts using similar doxycycline-inducible OSKM expression system (*39*).

Developing Neurons

For isolation of neurons at different developmental stages, we used a double-reporter mouse line *Btg2*^{RFP}/*Tubb3*^{GFP}, in which the proliferating progenitors are double negative (RFP−/GFP−), newborn neurons are double positive (RFP+/GFP+), and the cells positive for RFP but negative for GFP are (RFP+/GFP−) are the proliferating progenitors that were not used in this study. Lateral cortices dissected from E14.5 murine embryos were dissociated using a papain-based neural dissociation kit (Miltenyi Biotech, Germany) and the cell populations of interest were separated based on the RFP/GFP expression using FACS as described in detail elsewhere (*40*). The three types of sorted cells were then subjected to RNA sequencing (*40*) and mechanical characterization.

Mouse embryonic fibroblasts

Previously established, immortalized WT and CAV1KO mouse embryonic fibroblasts derived from WT and CAV1KO littermate C57BL/9 mice (*90*) were used in this study. Cells were cultured in DMEM medium (11960044, Gibco), supplemented with 10% fetal bovine serum (10270106, Gibco), 2 mM glutamine (25030081, Gibco), 100 U ml^{-1}/100 µg ml^{-1} penicillin/streptomycin (15070063, Gibco), at 37°C and 5% CO_2. Sub-confluent cells were collected for mechanical measurements by trypsinization (25200056, Gibco).

MCF10A ER-Src cell line

The MCF10A ER-Src cells were a kind gift from K. Struhl (Harvard Medical School, MA, USA). ER-Src is a fusion of the v-Src (viral non-receptor tyrosine kinase) with the ligand-binding domain of the estrogen receptor, that can be induced by cell treatment with tamoxifen (TAM)

(*70*). Cells were grown at 37°C under 5% CO_2 in DMEM/F12 medium (11039047, Gibco), supplemented with 5% charcoal (C6241, Sigma-Aldrich)-stripped horse serum (16050122, Gibco), 20 ng ml^{-1} EGF (AF-100-15, Peprotech), 10 mg ml^{-1} insulin (I9278, Sigma-Aldrich), 0.5 mg ml^{-1} hydrocortisone (H0888, Sigma-Aldrich), 100 ng ml^{-1} cholera toxin (C8052, Sigma-Aldrich), and 100 U ml^{-1}/100 µg ml^{-1} penicillin/streptomycin (15070063, Gibco). To induce the Src expression cells were plated at 50% confluency, and after allowing to adhere for 24 h, treated with 1 µM 4OH-TAM (H7904, Sigma-Aldrich) or with identical volume of ethanol as a control. Cells were characterized in adherent state using AFM at timepoints specified in the text.

MECHANICAL MEASUREMENTS

Mechanical characterization of cells using RT-DC

RT-DC measurements for mechanical characterization of cells were performed at room temperature according to previously established procedures (*91*). In brief, cells were harvested by trypsinization (adherent cells) and/or centrifugation at 400 g for 3–5 min, and suspended in a measurement buffer (MB). MB (osmolarity 310–315 mOsm kg^{-1}, pH 7.4) was based on phosphate buffered saline without Mg^{2+} and Ca^{2+} and contained 0.5% or 0.6% (w/w) methylcellulose (4000 cPs, Alfa Aesar, Germany) for increased viscosity. Cells were introduced into a microfluidic chip using a syringe pump (NemeSys, Cetoni, Germany), and focused into a 300-µm long channel constriction (with a square cross-section of 20 × 20 or 30 × 30 µm) by sheath flow infused at a flow rate three times as high as that of the cell suspension. The imaging was performed at the end of the channel constriction (Fig. S1B) at 2,000 frames s^{-1}. The cell area and deformation were derived from the fitted cell contours in real-time by the acquisition software (*ShapeIn2*; Zellmechanik Dresden, Germany). Apparent Young's modulus values were assigned to each cell based on its area and deformation under given experimental conditions (flow rate, channel size, viscosity of the medium, temperature) using a look-up table obtained through numerical simulations of an elastic solid(*33*) with the aid of *ShapeOut* (*ShapeOut 1.0.1*; https://github.com/ZELLMECHANIK-DRESDEN/ShapeOut;

Zellmechanik Dresden). The events were filtered for area ratio (the ratio between the area enclosed by the convex hull of the cell contour and the raw area enclosed by the contour) to discard incomplete contours or cells with rough surface, and for cell area and aspect ratio to discard derbies and doublets. Experimental details (channel sizes, flow rates, measurement buffers) and gates used for filtration in respective datasets are listed in Table S5.

Mechanical characterization of cells using AFM

For AFM measurements, cells were seeded on glass bottom dishes (FluoroDish; FD35100, WPI, FL, USA) at least one day in advance. Mechanical characterization was performed on adherent cells in a sub-confluent culture in CO_2-independent medium (18045054, Gibco) at 37°C (temperature was maintained by a petri dish heater, JPK Instruments, Germany). AFM measurements on TGBC and ECC4 cell lines were conducted on a *Nanowizard 4* (JPK Instruments). Tip-less cantilevers (PNP-TR-TL, nominal spring constant $k = 0.08\ \mathrm{N\ m^{-1}}$, Nanoworld, Switzerland) decorated a polystyrene bead of 5-µm diameter (PS-R-5.0, microParticles, Germany) each were used as the indenters. The cantilever spring constants were measured prior to each experiment using the thermal noise method implemented in the *JPK SPM* software (JPK Instruments). For each cell three indentation curves were recorded with a piezo extension speed of $5\ \mu\mathrm{m\ s^{-1}}$ to a maximum set force of 2 nN. For the microrheology analysis, the cantilever was lowered using a piezo extension speed of $5\ \mu\mathrm{m\ s^{-1}}$ until a force set point of 1 nN was reached, corresponding to an approximate indentation depth δ_0 of 1 µm. The lowered cantilever was then oscillated by a sinusoidal motion of the piezo elements at an amplitude of 10 nm for a period of 10 cycles. The oscillations were performed sequentially at different frequencies in the range of 3–200 Hz. Indentation experiments on MCF10A ER-Src cells were conducted as described above, except different tip-less cantilevers (Arrow TL1, nominal spring constant $k = 0.35–0.45\ \mathrm{N\ m^{-1}}$, Nanoworld) with a 5-µm bead glued at the end were used as the indenter.

AFM indentation data analysis

Recorded force-distance curves were converted into force-indentation curves and fitted in JPK data processing software (*JPK DP*, JPK Instruments) using Sneddon's modification of the Hertz model for a spherical indenter (*92*):

$$F = \frac{E}{1-\upsilon^2}\left(\frac{a^2+r^2}{2}\ln\frac{r+a}{r-a} - ar\right), \quad (3)$$

With

$$\delta = \frac{a}{2}\ln\frac{r+a}{r-a}, \quad (4)$$

where F denotes the indentation force, E the elastic modulus, υ the Poisson's ratio, a the radius of the projected contact area formed between the sample and the indenter, r the radius of the indenter, and δ the indentation depth. Poisson ratio was set to 0.5.

AFM microrheology data analysis

The force and indentation signals from oscillatory measurements were fitted using a sinusoidal function to extract the amplitude and phase angle of each signal. Data were analyzed analogously to the procedure described by Alcaraz et al. (*93*) but for a spherical not a pyramidal indenter. Briefly, the method relies on the linearization of the Hertz model for a spherical indenter due to small oscillations by using the first term of the Taylor expansion and subsequent transformation to the frequency domain:

$$F(\omega) = 2\frac{E^*(\omega)}{(1-\nu^2)}\sqrt{R\delta_0}\delta(\omega), \quad (5)$$

where $F(\omega)$ and $\delta(\omega)$ are the force and indentation signals in the frequency domain, respectively, $E^*(\omega)$ is the complex Young's modulus, ν is the Poisson's ratio assumed to be 0.5, R is the radius of the indenter and ω is the angular frequency. The complex shear modulus $G^*(\omega)$ can be written using $G^*(\omega) = \frac{E^*(\omega)}{2(1+\nu)}$:

$$G^*(\omega) = G'(\omega) + iG''(\omega) = \frac{(1-\nu)}{4\sqrt{R\delta_0}} \frac{F(\omega)}{\delta(\omega)}, \quad (6)$$

where $G'(\omega)$ is the storage modulus and $G''(\omega)$ is the loss modulus. The ratio of the force $F(\omega)$ and indentation $\delta(\omega)$ is calculated from the measured amplitudes $A^F(\omega)$ and $A^\delta(\omega)$ and the phase shifts $\theta^F(\omega)$ and $\theta^\delta(\omega)$ of the oscillatory signals (*94*):

$$\frac{F(\omega)}{\delta(\omega)} = \frac{A^F(\omega)}{A^\delta(\omega)} e^{i\left(\theta^F(\omega) - \theta^\delta(\omega)\right)}, \quad (7)$$

where the difference of the phase shifts $\left(\theta^F(\omega) - \theta^\delta(\omega)\right)$ is in the range of 0° (elastic solid) and 90° (viscous fluid). Furthermore, the hydrodynamic drag contribution on the cantilever oscillation was estimated and subtracted from the complex shear modulus as previously described (*95*):

$$G^*(\omega) = \frac{(1-\nu)}{4\sqrt{R\delta_0}} \left[\frac{F(\omega)}{\delta(\omega)} - i\omega b(0)\right], \quad (8)$$

where $b(h)$ is the hydrodynamic drag coefficient function measured from non-contact oscillations of the cantilever at different distances h from the sample, and $b(0)$ is the extrapolation to distance 0 from the sample. For PNP-TR-TL cantilevers, the hydrodynamic drag coefficient was estimated to be $b(0) = 5.28\ \mu\mathrm{N\ s\ m^{-1}}$.

PERTURBATION EXPERIMENTS

CAV1 knock-down

For RNAi experiments, cells were transfected using RNAiMax reagent (13778030, Thermo Fisher Scientific) and a reverse transfection protocol. Per transfection, 200 ng of esiRNA (Eupheria Biotech, Germany) or 300 ng of ON-TARGETplus siRNA (Dharmacon, CO, USA) and 2 µl RNAiMax were prepared in OptiMEM (31985062, Gibco) according to the manufacturer's instructions and pipetted onto 12-well plates (see Table S6 for full list of siRNAs used). Cells in 1 ml of culture medium were plated on top of the transfection mix at a density allowing for sub-confluent growth within the experimental timeframe. 72 h post transfection, cells were collected for the mechanical characterization and Western blot analysis.

Plasmid for CAV1 overexpression

The cDNA of *CAV1* was amplified by PCR, introducing NheI and XhoI restriction sites in the flanking regions. The PCR product was then cloned into the pCGIT destination vector (a kind gift from P. Serup, University of Copenhagen, Denmark) under the CAG promoter and with dTomato fluorescent marker under internal ribosomal entry site (IRES) downstream of CAV1.

Transient CAV1 overexpression in ECC4 and TGBC cells

ECC4 and TGBC cells were transiently transfected with the CAV1 overexpression plasmid by electroporation (Neon Transfection System, MPK5000, Thermo Fisher Scientific). Per transfection 0.3×10^6 ECC4 cells, or 0.2×10^6 TGBC cells were mixed with 1 µg of plasmid DNA in PBS. Electroporation was conducted using 10 µl Neon tips (MPK1096, Thermo Fisher Scientific) and a program of two pulses of 1050 V and 30 ms duration each. Electroporated cells were transferred to 500 µl of fresh culture medium in a 24-well plate. The cells were collected for mechanical characterization and Western blot analysis 72 h post transfection. To identify fluorescent cells during mechanical characterization, the combined real-time fluorescence and deformability cytometry (RT-FDC) (*20*) setup was used, and the maximum intensity of the fluorescence signal from channel 2 (excitation 561 nm, 10% laser power; collection 700/75) was utilized for gating.

Transient CAV1 overexpression in MCF10A-ER-src cells

MCF10A-ER-src cells were transiently transfected with the CAV1 overexpressing plasmid using Effectene transfection reagent (301425, Qiagen). One day before transfection, cells were seeded on glass bottom 35-mm dishes (FluoroDish; FD35100, WPI, FL, USA) at a density of 20,000 cells per well. Transfection was performed according to the manufacturer's instruction using 75 µl EC buffer, 0.6 µg plasmid DNA, 4.8 µl Enhancer and 6 µl Effectene reagent per well. 24 h post transfection cells were induced with 1 µM TAM. Mechanical analysis was performed after additional 72 h of culture.

Western blotting

For Western blot analysis of carcinoma and MCF10A-ER-Src cell lines, cell pellets were collected in parallel with mechanical measurements and lysed using ice-cold RIPA buffer (89900, ThermoFisher Scientific) supplemented with protease/phosphatase inhibitor cocktail (78441, ThermoFisher Scientific) and benzonase (E1014, Sigma-Aldrich). The lysates were cleared at 4°C by 10-minute sonication followed by 10-minute centrifugation at 16,900 g. Obtained supernatants were mixed with Laemmli buffer (final concertation: 62.5 mM Tris-HCl (pH 6.8), 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue), boiled (5 min at 95°C), and separated by SDS-PAGE electrophoresis on 4–20% gradient gels (Mini-PROTEAN TGX Precast Gels; 4561093, Biorad, CA, USA) in MOPS SDS Running buffer (B0001, ThermoFisher Scientific). After transferring the proteins onto a PVDF membrane (Merck Millipore), the membranes were blocked in TBS-T (20 mM Tris, 137 mM NaCl, 0.1% Tween) containing 5% w/v skimmed milk powder (T145.1, Carl Roth, Germany) for 40 minutes. Next, membranes were incubated with the primary anti-Cav1 (1:1000; D46G3; #3267, Cell Signaling Technology, MA, USA) and anti-GAPDH (1:5000; ab9485, Abcam, UK) antibodies at 4°C overnight in 5% milk/TBS-T, washed, and incubated with anti-rabbit HRP-conjugated secondary antibody (1:4000; ab97069, Abcam). Chemiluminescence detection was performed using Pierce Enhanced Chemi-Luminescence (ECL) substrate (32109, ThermoFisher Scientific) and ECL films (GE28-9068-37, Merck Millipore). Films were developed in an OptiMax X-ray film processor (KODAK, NY, USA). Quantitative analysis was performed on scanned films using the gel analysis tool in *JmageJ* version 2.0.0-rc-69/1.52p (https://imagej.nih.gov/). For western blot analysis of MEFs the same anti-Cav1 antibody (1:1000; D46G3; #3267, Cell Signaling) was used, and anti-tubulin antibody (1:2000; DM1A; #3873, Cell Signaling) was used as a loading control. Goat anti-mouse 680 and goat anti-rabbit 800 (1:2000; A28183 and A32735, ThermoFisher Scientific) antibodies were used for secondary detection. Membranes were scanned with the Odyssey imaging system (LI-COR Biosciences, NE, USA).

COMPUTATIONAL ANALYSIS

Transcriptomic datasets

Transcriptomic datasets were retrieved from online databases (Gene Expression Omnibus, GEO and DNA Data Bank of Japan, DDBJ) with accession numbers listed in Table 1. Overview of experimental detail for RNA profiling procedures and data analysis in individual datasets is presented in Table S7. The IDs of samples used in respective categories in each dataset are listed in Table S8. In case of multiple entries for the same gene in a given transcriptomic dataset, the expression values were averaged, so that only one entry per gene and sample was available.

PC-corr analysis

Before performing the PC-corr analysis, the glioblastoma and iPSC datasets were intersected and normalized by taking the log10 (glioblastoma dataset) or z-score (iPSC dataset) of the subset of 9,452 overlapping genes. The PC-corr analysis was conducted on individual datasets as described in detail elsewhere (*29*). In brief, PCA was performed using *svd* function in *MATLAB* (*R2020a*, MathWorks, MA, USA) on normalized datasets. The original PC loadings from the component providing good separation of sample categories (PC1 for both analyzed datasets) were processed in a two-step procedure including the normalization and scaling. The processing of the PC loadings is performed to adjust the distribution of the loadings to the range of Pearson's correlation values [–1,1], so that they are comparable when computing the *PC-corr* value. The normalization was performed using a custom function developed previously (*29*) of the following formula:

$$V_i^* = \mathrm{sgn}\left(V_i^0\right) \log_{10}\left(1 + \frac{|V_i^0|}{\langle |V^0| \rangle}\right), \tag{9}$$

where V_i^* denotes the normalized loading corresponding to the i-th feauture, V_i^0 the original loading corresponding to the i-th feauture, and $\langle |V^0| \rangle$ the average of all absolute loadings of the vector V^0.

The normalized loadings were then scaled to fall on the interval [–1,1] using a previously developed custom function (*29*):

$$V_i = \mathrm{sgn}(V_i^*)\ \frac{|V_i^*| - \min(|V^*|)}{\max(|V^*|) - \min(|V^*|)}, \tag{10}$$

where V_i denotes the processed loading corresponding to the *i*-th feature, and $|V^*|$ the vector containing absolute values of all normalized loadings.

The *PC-corr* values for each pair of features were computed according to Equation 1. The PC-corr results of the glioblastoma and iPSC datasets were combined as described in the results section. Gene pairs showing different $PC\text{-}corr$ signs were masked by setting the $PC\text{-}corr^{comb}$ to zero. The genes and edges comprising the network were obtained via thresholding strategies described in the main text. The network was visualized using *cytoscape* (*cytoscape 3.8.0*; https://cytoscape.org/) (*96*).

Combinatorial marker

To compute a combinatorial marker associated with a gene functional network module composed of *n* genes, we use the following three-step procedure:

(1) *Dataset normalization*: To scale the features to a comparable range and reduce the dominant influence of highly expressed genes, each dataset is normalized. Possible normalization approaches include logarithm normalization ($x = \log(x+1)$ and *z*-score normalization. Since both normalization approaches lead to comparable results, we decided to proceed with the logarithm normalization because it is one of the most widely adopted in computational genomics for combinatorial markers (*97*).

(2) *Direction alignment*: The direction of the gene expression change between different samples is analyzed and, if necessary, aligned. The pairwise Pearson's correlation of the *n* genes from the dataset used for inference of the combinatorial marker is computed. If all pairs of genes are positively correlated between each other there is no need of direction alignment — this was the specific case in our study. Otherwise, the directions of genes whose correlation with the reference gene is negative need to be aligned before the compression.

The reference gene for the direction alignment is the gene with the highest average pairwise Pearson's correlation with the other $n-1$ genes in the functional module. The alignment is performed by subtracting the mean value of the normalized expression across samples, $\bar{g}$, from the normalized expression of the given gene, g, inverting its trend using the multiplication by −1, and finally adding again the mean value to regain the original expression level:

$$\mathrm{align}(g) = -(g - \bar{g}) + \bar{g} = 2\bar{g} - g. \tag{11}$$

Once defined, the aligned values should be used for any further validation analysis, including the computation of the JVT. The alignment step is necessary to make sure that the information contained in the anticorrelated genes does not annihilate each other during the compression into the combinatorial marker. Below, an example is provided to illustrate this issue.

(3) *Compression*: To perform compression and obtain the combinatorial marker g_{comb} we employ one of the most employed compression operators in computational genomics (*97*), the mean operator:

$$g_{comb} = \frac{1}{n}\sum_{i=1}^{n} g_i \,, \tag{12}$$

where g_i indicates the normalized and aligned expression value of the *i*-th gene of the functional module from which the combinatorial marker is derived.

To illustrate the importance of the alignment, let us consider a simple example of two anticorrelated genes in four samples: g_1 = [1 1 3 3] and g_2 = [3 3 1 1]. When the compression is performed without alignment, following values of the combinatorial marker are obtained: $g_{comb} = \frac{g_1+g_2}{2}$ = [2 2 2 2]. The so-obtained combinatorial marker is non-discriminative, even though the individual genes are. On the contrary, if the alignment function is applied prior to compression: $g_{comb} = \frac{g_1 + \mathrm{align}(g_2)}{2} = \frac{[1\ 1\ 3\ 3] + [\,4 - [\,3\ 3\ 1\ 1]\,]}{2} = \frac{[1\ 1\ 3\ 3] + [\,1\ 1\ 3\ 3]\,]}{2} = [1\ 1\ 3\ 3]$, the original discriminative information is conserved in the combined marker.

Joint-view trustworthiness (JVT)

The single-view trustworthiness measure was recently introduced by us in studies on pattern recognition to assess the extent to which the geometrical discrimination of samples of a dataset might emerge at random along a dimension of embedding in a geometrical space (*74*, *98*). In brief, the single-view trustworthiness measure is an empirical *p* value computed from a null model distribution obtained by a resampling technique, which randomly shuffles the labels of the samples and computes what is the probability to generate at random a matching between sample labels and sample geometrical location that offers a discrimination that is equal or larger than the one tested. The obtained *p* value assesses whether the visualized and measured sample discrimination along a dimension of a geometrical space is significant (because rare to appear at random) or no significant (because frequent to appear at random). This is particularly useful to assess the trustworthiness of a discriminative result when the number of samples for each class is small or when it is unbalanced, as is the case for some datasets in our study. To assess the trustworthiness of a marker's discrimination performance jointly on many datasets, we introduce a joint-view extension to this method which we refer to as the joint-view trustworthiness (JVT).

To ensure that the proposed markers have a joint multiview discrimination that is rare to obtain by chance, JVT samples markers at random from the data and compares their performance to the one of predicted targets according to the following procedure:

(1) *Data preparation*: Collect datasets that support (positive hypothesis: for instance, discriminative presence of a cell mechanic phenotype: soft/stiff) or not support (negative hypothesis: for instance, discriminative absence of a cell mechanic phenotype: soft/stiff) your hypothesis, and make sure that you consider for all of them only the features (genes) that are common to each dataset in you study.

(2) *Data normalization*: Perform only when computing combinatorial marker (see point (1) in the combinatorial marker section above).

(3) *Null model distribution sampling and p-value estimation*:

(3.1) *Single marker test:* Sample at random a gene and extract its expression from each dataset, compute its joint multiview discrimination performance as the minimum performance measure (we adopted the AUC-ROC because it is one of the most used in classification assessment, but any classification performance measure can be employed) across the datasets. Repeat this procedure sampling at random for *T* times (in our study we used *T* = 10,000) a gene from the datasets and computing its minimum classification performance measure across the datasets. The ensemble of the *T* minimum classification performance measures can be used to draw an empirical distribution that forms the null model. The *p* value of the tested marker is computed counting the proportion of genes that within the *T* samplings have a minimum classification performance that is equal or larger than that for the tested marker. Please note that here we compute the joint multiview discrimination performance using the minimum performance across the datasets because we pursue a conservative estimation. Other operators, such as mean, median, or mode can be employed instead of the minimum operator to make the JVT estimation less conservative.

(3.2) *Combinatorial marker test*: Given a combinatorial marker of *m* genes, sample at random *m* genes and extract their expressions from each dataset, compute the combinatorial marker (apply the same compression formula of the tested combinatorial marker) joint multiview discrimination performance as the minimum performance measure (we adopted the AUC-ROC, see point (3.1) above for details) across the datasets. Repeat this procedure *T* times (in our study we used *T* = 10,000). The *p* value is computed as for the single marker test (see point (3.1) above).

The JVT pseudocode and time complexity analysis are provided in Supplementary Material. In brief, the overall complexity of JVT considering a scenario like in our study is O(Z+T), i.e., JVT is linear in Z (number of common genes in the datasets) and T (number of samplings).

The JVT code (in MATLAB, R and Python) and datasets to replicate the results in Table 4 of this study are available at: https://github.com/biomedical-cybernetics/Joint-View-trustworthiness-JVT.

Statistical analysis

The RT-DC datasets were compared using generalized linear mixed effects models with the aid of *ShapeOut* (*ShapeOut 1.0.1*; https://github.com/ZELLMECHANIK-DRESDEN/ShapeOut; Zellmechanik Dresden) as described in detail elsewhere (*99*). AFM datasets were compared using two-sided Wilcoxon rank sum test in *MATLAB* (*R2020a*, MathWorks). Western blot results were compared using a two-sided two-sample t-test in *MATLAB* (*R2020a*, MathWorks).

Acknowledgments

We thank Isabel Richter and Christine Schweitzer for technical assistance, Miguel Sanchez (CNIC, Spain) and Konstantinos Anastasiadis (TU Dresden, Germany) for helpful discussions, Len R. Stephens (Babraham Institute, UK) for provision of MCF10A PIK3CA cells, and Kevin Struhl (Harvard Medical School, MA, USA) for provision of MCF10A ER-Src cells. We further thank the Microstructure Facility at the Center for Molecular and Cellular Bioengineering (CMCB) at the Technische Universität Dresden (in part funded by the State of Saxony and the European Regional Development Fund) for hosting the chip fabrication.

The authors acknowledge the following funding:

Alexander von Humboldt-Stiftung, Alexander von Humboldt Professorship (JG)

European Commission, ERC Starting Grant "LightTouch" #282060 (JG)

Marie Sklodowska-Curie Actions under the European Union's Horizon 2020 research and innovation programme, BIOPOL ITN, #641639 (MADP, JG)

Deutsche Forschungsgemeinschaft, #GU 612/5-1 and #399422891 (JG)

Comunidad Autónoma de Madrid, Tec4Bio-CM, #S2018/NMT-4443 (MADP)

Fundació La Marató de TV3, #201936-30-31 (MADP)

Mildred Scheel Early Career Center Dresden (MSNZ) funded by the German Cancer Aid (Deutsche Krebshilfe) (AT)

Author contributions

JG and CVC conceived the project. YG under supervision of CVC performed the computational analysis on transcriptomics data. CVC invented the joint-view trustworthiness (JVT) theory and algorithm. SSA, under the supervision of CVC, wrote the JVT pseudo-code and the time complexity analysis. CVC coded the MATLAB function and SSA coded the R and Python functions. All codes were benchmarked by SSA, YG and CVC. MU, MW, MH, MK, and NT performed the measurements of mechanical states. MU and MW performed the genetic manipulation experiments. SA and AT provided methodological support with AFM measurements and data analysis. JD and OF were involved in MCF10A PIK3CA experiments.

MD under supervision of FC isolated the developing neurons. FL under supervision of MADP. performed MEF experiments. MU and MW analyzed the experimental data. MU visualized the data and prepared figures. MU, with the support of JG, CVC, and YG prepared the initial version of the manuscript. All authors revised and edited the manuscript. JG, CVC, MADP and AT acquired funding.

Competing interests

S.A., M.K., and J.G. are co-founders and shareholders of the company Rivercyte GmbH that is commercializing deformability cytometry technology. The remaining authors declare no competing interests.

Data and materials availability

All data needed to evaluate the conclusions in the paper are present in the manuscript or available on public repositories as detailed below. The transcriptomic data used in this study were obtained from public repositories, their accession numbers are listed in Table 1. The mechanical characterization data are available as a collection on figshare: https://doi.org/10.6084/m9.figshare.c.5399826**.** The MATLAB code for performing the PC-corr analysis was based on the code deposited alongside a previous publication (*29*), accessible on GitHub: https://github.com/biomedical-cybernetics/PC-corr_net. The JVT code (in MATLAB, R and Python) and datasets for replicating the results presented in Table 4 are available on GitHub: https://github.com/biomedical-cybernetics/Joint-View-trustworthiness-JVT and figshare: https://doi.org/10.6084/m9. figshare.20123159.

Supplementary Material

Supplementary Material file is provided alongside the manuscript and includes:

- Figs. S1 to S7
- Tables S1 to S8
- Joint-view trustworthiness (JVT) pseudocode and computational complexity analysis

References

1. J. Guck, E. R. Chilvers, Mechanics meets medicine. *Sci. Transl. Med.* **5**, 3–6 (2013).
2. Y. Nematbakhsh, C. T. Lim, Cell biomechanics and its applications in human disease diagnosis. *Acta Mech. Sin.* **31**, 268–273 (2015).
3. J. Guck, Some thoughts on the future of cell mechanics. *Biophys. Rev.* **11**, 667–670 (2019).
4. T. Lecuit, P. F. Lenne, Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nat. Rev. Mol. Cell Biol.* **8**, 633–644 (2007).
5. E. Hannezo, C. P. Heisenberg, Mechanochemical Feedback Loops in Development and Disease. *Cell*. **178**, 12–25 (2019).
6. S. Suresh, Biomechanics and biophysics of cancer cells. *Acta Mater.* **55**, 3989–4014 (2007).
7. V. Gensbittel, M. Kräter, S. Harlepp, I. Busnelli, J. Guck, J. G. Goetz, Mechanical Adaptability of Tumor Cells in Metastasis. *Dev. Cell*. **56**, 164–179 (2021).
8. D. A. Fletcher, R. D. Mullins, Cell mechanics and the cytoskeleton. *Nature*. **463**, 485–492 (2010).
9. P. Chugh, E. K. Paluch, The actin cortex at a glance. *J. Cell Sci.* **131**, 1–9 (2018).
10. M. Kelkar, P. Bohec, G. Charras, Mechanics of the cellular actin cortex: From signalling to shape change. *Curr. Opin. Cell Biol.* **66**, 69–78 (2020).
11. K. Seltmann, A. W. Fritsch, J. A. Käs, T. M. Magin, Keratins significantly contribute to cell stiffness and impact invasive behavior. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18507–18512 (2013).
12. A. E. Patteson, R. J. Carroll, D. V Iwamoto, P. A. Janmey, The vimentin cytoskeleton: When polymer physics meets cell biology. *Phys. Biol.* **18**, 11001 (2020).
13. H. Kubitschke, J. Schnauss, K. D. Nnetu, E. Warmt, R. Stange, J. Kaes, Actin and microtubule networks contribute differently to cell response for small and large strains. *New J. Phys.* **19** (2017), doi:10.1088/1367-2630/aa7658.
14. Y. C. Chang, P. Nalbant, J. Birkenfeld, Z. F. Chang, G. M. Bokoch, GEF-H1 couples nocodazole-induced microtubule disassembly to cell contractility via RhoA. *Mol. Biol. Cell*. **19**, 2147–2153 (2008).
15. E. H. Zhou, X. Trepat, C. Y. Park, G. Lenormand, M. N. Oliver, S. M. Mijailovich, C. Hardin, D. A. Weitz, J. P. Butler, J. J. Fredberg, Universal behavior of the osmotically compressed cell and its analogy to the colloidal glass transition. *PNAS*. **106**, 10632–10637 (2009).
16. M. Guo, A. F. Pegoraro, A. Mao, E. H. Zhou, P. R. Arany, Y. Han, D. T. Burnette, M. H. Jensen, K. E. Kasza, J. R. Moore, F. C. Mackintosh, J. J. Fredberg, D. J. Mooney, J. Lippincott-Schwartz, D. A. Weitz, Cell volume change through water efflux impacts cell stiffness and stem cell fate. *PNAS*. **c**, 201705179 (2017).
17. N. Caille, O. Thoumine, Y. Tardy, J. J. Meister, Contribution of the nucleus to the mechanical properties of endothelial cells. *J. Biomech.* **35**, 177–187 (2002).
18. P. Chugh, A. G. Clark, M. B. Smith, D. A. D. Cassani, K. Dierkes, A. Ragab, P. P. Roux, G. Charras,

G. Salbreux, E. K. Paluch, Actin cortex architecture regulates cell surface tension. *Nat. Cell Biol.* **19**, 689–697 (2017).

19. Y. Toyoda, C. J. Cattin, M. P. Stewart, I. Poser, M. Theis, T. V. Kurzchalia, F. Buchholz, A. A. Hyman, D. J. Müller, Genome-scale single-cell mechanical phenotyping reveals disease-related genes involved in mitotic rounding. *Nat. Commun.* **8**, 1–10 (2017).

20. P. Rosendahl, K. Plak, A. Jacobi, M. Kraeter, N. Toepfner, O. Otto, C. Herold, M. Winzi, M. Herbig, Y. Ge, S. Girardo, K. Wagner, B. Baum, J. Guck, M. Kräter, N. Töpfner, O. Otto, C. Herold, M. Winzi, Y. Ge, S. Girardo, K. Wagner, J. Guck, M. Kraeter, N. Toepfner, O. Otto, C. Herold, M. Winzi, M. Herbig, Y. Ge, S. Girardo, K. Wagner, B. Baum, J. Guck, Real-time fluorescence and deformability cytometry. *Nat. Methods*. **15**, 355 (2018).

21. J. Wang, D. Lü, D. Mao, M. Long, Mechanomics: An emerging field between biology and biomechanics. *Protein Cell*. **5**, 518–531 (2014).

22. V. D. L. Putra, M. J. Song, S. McBride-Gagyi, H. Chang, K. Poole, R. Whan, D. Dean, V. Sansalone, M. L. Knothe Tate, Mechanomics Approaches to Understand Cell Behavior in Context of Tissue Neogenesis, During Prenatal Development and Postnatal Healing. *Front. Cell Dev. Biol.* **7**, 1–8 (2020).

23. F. Zhang, J. Wang, D. Lü, L. Zheng, B. Shangguan, Y. Gao, Y. Wu, M. Long, Mechanomics analysis of hESCs under combined mechanical shear, stretch, and compression. *Biomech. Model. Mechanobiol.* (2020), doi:10.1007/s10237-020-01378-5.

24. M. Lang, Lighting up the mechanome. *Front. Eng. Reports leading-edge Eng. from 2007 Symp.* (2008).

25. J. J. W. A. Van Loon, Mechanomics and physicomics in gravisensing. *Microgravity Sci. Technol.* **21**, 159–167 (2009).

26. M. J. Song, S. M. Brady-Kalnay, S. H. McBride, P. Phillips-Mason, D. Dean, M. L. K. Tate, Mapping the mechanome of live stem cells using a novel method to measure local strain fields in situ at the fluid-cell interface. *PLoS One*. **7**, 1–12 (2012).

27. M. J. Song, D. Dean, M. L. Knothe Tate, Mechanical modulation of nascent stem cell lineage commitment in tissue engineering scaffolds. *Biomaterials*. **34**, 5766–5775 (2013).

28. H. Wang, H. Zhang, B. Da, D. Lu, R. Tamura, K. Goto, I. Watanabe, D. Fujita, N. Hanagata, J. Kano, T. Nakagawa, M. Noguchi, Mechanomics Biomarker for Cancer Cells Unidenti fi able through Morphology and Elastic Modulus (2021), doi:10.1021/acs.nanolett.1c00003.

29. S. Ciucci, Y. Ge, C. Durán, A. Palladini, V. Jiménez-Jiménez, L. M. Martínez-Sánchez, Y. Wang, S. Sales, A. Shevchenko, S. W. Poser, M. Herbig, O. Otto, A. Androutsellis-Theotokis, J. Guck, M. J. Gerl, C. V. Cannistraci, Enlightening discriminative network functional modules behind principal component analysis separation in differential-omic science studies. *Sci. Rep.* **7**, 1–24 (2017).

30. S. W. Poser, O. Otto, C. Arps-Forker, Y. Ge, M. Herbig, C. Andree, K. Gruetzmann, M. F. Adasme, S. Stodolak, P. Nikolakopoulou, D. M. Park, A. Mcintyre, M. Lesche, A. Dahl, P. Lennig, S. R. Bornstein, E. Schroeck, B. Klink, R. R. Leker, M. Bickle, G. P. Chrousos, M. Schroeder, C. V. Cannistraci, J. Guck, A. Androutsellis-Theotokis, Controlling distinct signaling states in cultured cancer cells provides a new platform for drug discovery. *FASEB J.* **33**, 9235–9249 (2019).

31. M. Urbanska, M. Winzi, K. Neumann, S. Abuhattum, P. Rosendahl, P. Müller, A. Taubenberger, K. Anastassiadis, J. Guck, Single-cell mechanical phenotype is an intrinsic marker of reprogramming and differentiation along the mouse neural lineage. *Development*. **144**, 4313–4321 (2017).

32. O. Otto, P. Rosendahl, A. Mietke, S. Golfier, C. Herold, D. Klaue, S. Girardo, S. Pagliara, A. Ekpenyong, A. Jacobi, M. Wobus, N. Töpfner, U. F. Keyser, J. Mansfeld, E. Fischer-Friedrich, J. Guck, Real-time deformability cytometry: on-the-fly cell mechanical phenotyping. *Nat. Methods*. **12**, 199–202 (2015).

33. M. Mokbel, D. Mokbel, A. Mietke, N. Träber, G. Salvatore, O. Otto, J. Guck, S. Aland, Numerical Simulation of Real-Time Deformability Cytometry To Extract Cell Mechanical Properties. *ACS Biomater. Sci. Eng.* **3**, 2962–2973 (2017).

34. A. R. R. Forrest, H. Kawaji, M. Rehli, J. K. Baillie, M. J. L. De Hoon, V. Haberle, T. Lassmann, I. V. Kulakovskiy, M. Lizio, M. Itoh, R. Andersson, C. J. Mungall, T. F. Meehan, S. Schmeier, N. Bertin, M. Jørgensen, E. Dimont, E. Arner, C. Schmidl, U. Schaefer, Y. A. Medvedeva, C. Plessy, M. Vitezic, J. Severin, C. A. Semple, Y. Ishizu, R. S. Young, M. Francescatto, I. A. Altschuler, D. Albanese, G. M. Altschule, T. Arakawa, J. A. C. Archer, P. Arner, M. Babina, S. Rennie, P. J. Balwierz, A. G. Beckhouse, S. Pradhan-Bhatt, J. A. Blake, A. Blumenthal, B. Bodega, A. Bonetti, J. Briggs, F. Brombacher, A. M. Burroughs, A. Califano, C. V. Cannistraci, D. Carbajo, Y. Chen, M. Chierici, Y. Ciani, H. C. Clevers, E. Dalla, C. A. Davis, M. Detmar, A. D. Diehl, T. Dohi, F. Drabløs, A. S. B. Edge, M. Edinger, K. Ekwall, M. Endoh, H. Enomoto, M. Fagiolini, L. Fairbairn, H. Fang, M. C. Farach-Carson, G. J. Faulkner, A. V. Favorov, M. E. Fisher, M. C. Frith, R. Fujita, S. Fukuda, C. Furlanello, M. Furuno, J. I. Furusawa, T. B. Geijtenbeek, A. P. Gibson, T. Gingeras, D. Goldowitz, J. Gough, S. Guhl, R. Guler, S. Gustincich, T. J. Ha, M. Hamaguchi, M. Hara, M. Harbers, J. Harshbarger, A. Hasegawa, Y. Hasegawa, T. Hashimoto, M. Herlyn, K. J. Hitchens, S. J. H. Sui, O. M. Hofmann, I. Hoof, F. Hori, L. Huminiecki, K. Iida, T. Ikawa, B. R. Jankovic, H. Jia, A. Joshi, G. Jurman, B. Kaczkowski, C. Kai, K. Kaida, A. Kaiho, K. Kajiyama, M. Kanamori-Katayama, A. S. Kasianov, T. Kasukawa, S. Katayama, S. Kato, S. Kawaguchi, H. Kawamoto, Y. I. Kawamura, T. Kawashima, J. S. Kempfle, T. J. Kenna, J. Kere, L. M. Khachigian, T. Kitamura, S. P. Klinken, A. J. Knox, M. Kojima, S. Kojima, N. Kondo, H. Koseki, S. Koyasu, S. Krampitz, A. Kubosaki, A. T. Kwon, J. F. J. Laros, W. Lee, A. Lennartsson, K. Li, B. Lilje, L. Lipovich, A. Mackay-sim, R. I. Manabe, J. C. Mar, B. Marchand, A. Mathelier, N. Mejhert, A. Meynert, Y. Mizuno, D. A. L. De Morais, H. Morikawa, M. Morimoto, K. Moro, E. Motakis, H. Motohashi, C. L. Mummery, M. Murata, S. Nagao-Sato, Y. Nakachi, F. Nakahara, T. Nakamura, Y. Nakamura, K. Nakazato, E. Van Nimwegen, N. Ninomiya, H. Nishiyori, S. Noma, T. Nozaki, S. Ogishima, N. Ohkura, H. Ohmiya, H. Ohno, M. Ohshima, M. Okada-Hatakeyama, Y. Okazaki, V. Orlando, D. A. Ovchinnikov, A. Pain, R. Passier, M. Patrikakis, H. Persson, S. Piazza, J. G. D. Prendergast, O. J. L. Rackham, J. A. Ramilowski, M. Rashid, T. Ravasi, P. Rizzu, M. Roncador, S. Roy, M. B. Rye, E. Saijyo, A. Sajantila, A. Saka, S. Sakaguchi, M. Sakai, H. Sato, H. Satoh, S. Savvi, A. Saxena, C. Schneider, E. A. Schultes, G. G. Schulze-Tanzil, A. Schwegmann, T. Sengstag, G. Sheng, H. Shimoji, Y. Shimoni, J. W. Shin, C. Simon, D. Sugiyama, T. Sugiyama, M. Suzuki, N. Suzuki, R. K. Swoboda, P. A. C. 'T Hoen, M. Tagami, N. T. Tagami, J. Takai, H. Tanaka, H. Tatsukawa, Z. Tatum, M. Thompson, H. Toyoda, T. Toyoda, E. Valen, M. Van De Wetering, L. M. Van Den Berg, R. Verardo, D. Vijayan, I. E. Vorontsov, W. W. Wasserman, S. Watanabe, C. A. Wells, L. N. Winteringham, E. Wolvetang, E. J. Wood, Y. Yamaguchi, M. Yamamoto, M. Yoneda, Y. Yonekura, S. Yoshida, S. E. Zabierowski, P. G. Zhang, X. Zhao, S. Zucchelli, K. M. Summers, H. Suzuki, C. O. Daub, J. Kawai, P. Heutink, W. Hide, T. C. Freeman, B.

Lenhard, L. V. B. Bajic, M. S. Taylor, V. J. Makeev, A. Sandelin, D. A. Hume, P. Carninci, Y. Hayashizaki, A promoter-level mammalian expression atlas. *Nature*. **507**, 462–470 (2014).

35. J. Barretina, 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, L. A. Garraway, The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. **483**, 603–607 (2012).

36. M. Ghandi, F. W. Huang, J. Jané-Valbuena, G. V. Kryukov, C. C. Lo, E. R. McDonald, J. Barretina, E. T. Gelfand, C. M. Bielski, H. Li, K. Hu, A. Y. Andreev-Drakhlin, J. Kim, J. M. Hess, B. J. Haas, F. Aguet, B. A. Weir, M. V. Rothberg, B. R. Paolella, M. S. Lawrence, R. Akbani, Y. Lu, H. L. Tiv, P. C. Gokhale, A. de Weck, A. A. Mansour, C. Oh, J. Shih, K. Hadi, Y. Rosen, J. Bistline, K. Venkatesan, A. Reddy, D. Sonkin, M. Liu, J. Lehar, J. M. Korn, D. A. Porter, M. D. Jones, J. Golji, G. Caponigro, J. E. Taylor, C. M. Dunning, A. L. Creech, A. C. Warren, J. M. McFarland, M. Zamanighomi, A. Kauffmann, N. Stransky, M. Imielinski, Y. E. Maruvka, A. D. Cherniack, A. Tsherniak, F. Vazquez, J. D. Jaffe, A. A. Lane, D. M. Weinstock, C. M. Johannessen, M. P. Morrissey, F. Stegmeier, R. Schlegel, W. C. Hahn, G. Getz, G. B. Mills, J. S. Boehm, T. R. Golub, L. A. Garraway, W. R. Sellers, Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature*. **569**, 503–508 (2019).

37. C. Klijn, S. Durinck, E. W. Stawiski, P. M. Haverty, Z. Jiang, H. Liu, J. Degenhardt, O. Mayba, F. Gnad, J. Liu, G. Pau, J. Reeder, Y. Cao, K. Mukhyala, S. K. Selvaraj, M. Yu, G. J. Zynda, M. J. Brauer, T. D. Wu, R. C. Gentleman, G. Manning, R. L. Yauch, R. Bourgon, D. Stokoe, Z. Modrusan, R. M. Neve, F. J. De Sauvage, J. Settleman, S. Seshagiri, Z. Zhang, A comprehensive transcriptional portrait of human cancer cell lines. *Nat. Biotechnol.* **33**, 306–312 (2015).

38. V. Y. Kiselev, V. Juvin, M. Malek, N. Luscombe, P. Hawkins, N. Le Novère, L. Stephens, Perturbations of PIP3 signalling trigger a global remodelling of mRNA landscape and reveal a transcriptional feedback loop. *Nucleic Acids Res.* **43**, 9663–9679 (2015).

39. P. D. Tonge, A. J. Corso, C. Monetti, S. M. I. Hussein, M. C. Puri, I. P. Michael, M. Li, D.-S. Lee, J. C. Mar, N. Cloonan, D. L. Wood, M. E. Gauthier, O. Korn, J. L. Clancy, T. Preiss, S. M. Grimmond, J.-Y. Shin, J.-S. Seo, C. a. Wells, I. M. Rogers, A. Nagy, Divergent reprogramming routes lead to alternative stem-cell states. *Nature*. **516**, 192–197 (2014).

40. J. Aprea, S. Prenninger, M. Dori, T. Ghosh, L. S. Monasor, E. Wessendorf, S. Zocher, S. Massalini, D. Alexopoulou, M. Lesche, A. Dahl, M. Groszer, M. Hiller, F. Calegari, Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment. *EMBO J.* **32**, 3145–3160 (2013).

41. B. Brenner, L. H. Tang, D. S. Klimstra, D. P. Kelsen, Small-cell carcinomas of the gastrointestinal tract: A review. *J. Clin. Oncol.* **22**, 2730–2739 (2004).

42. G. P. Kalemkerian, W. Akerley, P. Bogner, H. Borghaei, L. Q. Chow, R. J. Downey, L. Gandhi, A. K. P. Ganti, R. Govindan, J. C. Grecula, J. Hayman, R. S. Heist, L. Horn, T. Jahan, M. Koczywas, B. W. Loo, R. E. Merritt, C. A. Moran, H. B. Niell, J. O'Malley, J. D. Patel, N. Ready, C. M. Rudin, C. C. Williams, K. Gregory, M. Hughes, Small cell lung cancer: Clinical practice guidelines in oncology.

JNCCN J. Natl. Compr. Cancer Netw. **11**, 78–98 (2013).

43. V. Juvin, M. Malek, K. E. Anderson, C. Dion, T. Chessa, C. Lecureuil, G. J. Ferguson, S. Cosulich, P. T. Hawkins, L. R. Stephens, Signaling via Class IA Phosphoinositide 3-Kinases (PI3K) in Human, Breast-Derived Cell Lines. *PLoS One*. **8** (2013), doi:10.1371/journal.pone.0075045.

44. A. G. Bader, S. Kang, P. K. Vogt, Cancer-specific mutations in PIK3CA are oncogenic in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 1475–1479 (2006).

45. S. Kang, A. G. Bader, P. K. Vogt, Phosphatidylinositol 3-kinase mutations identified in human cancer are oncogenic. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 802–807 (2005).

46. B. Sinha, D. Köster, R. Ruez, P. Gonnord, M. Bastiani, D. Abankwa, R. V. Stan, G. Butler-Browne, B. Vedie, L. Johannes, N. Morone, R. G. Parton, G. Raposo, P. Sens, C. Lamaze, P. Nassoy, Cells respond to mechanical stress by rapid disassembly of caveolae. *Cell*. **144**, 402–413 (2011).

47. R. G. Parton, M. A. Del Pozo, Caveolae as plasma membrane sensors, protectors and organizers. *Nat. Rev. Mol. Cell Biol.* **14**, 98–112 (2013).

48. F.-N. Lolo, N. Walani, E. Seemann, D. Zalvidea, D. M. Pavón, G. Cojoc, M. Zamai, C. Viaris de Lesegno, F. Martínez de Benito, M. Sánchez-Álvarez, J. J. Uriarte, A. Echarri, D. Jiménez-Carretero, J.-C. Escolano, S. A. Sánchez, V. R. Caiolfa, D. Navajas, X. Trepat, J. Guck, C. Lamaze, P. Roca-Cusachs, M. M. Kessels, B. Qualmann, M. Arroyo, M. A. del Pozo, M. A. Pozo, Caveolin-1 dolines form a distinct and rapid caveolae-independent mechanoadaptation system. *Nat. Cell Biol. 2022 251*. **25**, 120–133 (2023).

49. M. Johannessen, S. Møler, T. Hansen, U. Moens, M. Van Ghelue, The multifunctional roles of the four-and-a-half-LIM only protein FHL2. *Cell. Mol. Life Sci.* **63**, 268–284 (2006).

50. N. Nakazawa, A. R. Sathe, G. V. Shivashankar, M. P. Sheetz, Matrix mechanics controls FHL2 movement to the nucleus to activate p21 expression. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6813–E6822 (2016).

51. X. Sun, D. Y. Z. Phua, L. Axiotakis, M. A. Smith, E. Blankman, R. Gong, R. C. Cail, S. Espinosa de los Reyes, M. C. Beckerle, C. M. Waterman, G. M. Alushin, Mechanosensing through Direct Binding of Tensed F-Actin by LIM Domains. *Dev. Cell* (2020), doi:10.1016/j.devcel.2020.09.022.

52. J. D. Winkelman, C. A. Anderson, C. Suarez, D. R. Kovar, M. L. Gardel, Evolutionarily diverse LIM domain-containing proteins bind stressed actin filaments through a conserved mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 25532–25542 (2020).

53. L. Jin, F. Shen, M. Weinfeld, C. Sergi, Insulin Growth Factor Binding Protein 7 (IGFBP7)-Related Cancer and IGFBP3 and IGFBP7 Crosstalk. *Front. Oncol.* **10**, 1–14 (2020).

54. R. Liu, M. M. Hossain, X. Chen, J. P. Jin, Mechanoregulation of SM22α/Transgelin. *Biochemistry*. **56**, 5526–5538 (2017).

55. M. Dvorakova, R. Nenutil, P. Bouchal, Transgelins, cytoskeletal proteins implicated in different aspects of cancer development. *Expert Rev. Proteomics*. **11**, 149–165 (2014).

56. J. C. Adams, J. Lawler, The thrombospondins. *Cold Spring Harb. Perspect. Biol.* **3**, 1–29 (2011).

57. T. Huang, L. Sun, X. Yuan, H. Qiu, Thrombospondin-1 is a multifaceted player in tumor progression. *Oncotarget*. **8**, 84546–84558 (2017).

58. J. A. Hanley, B. J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. **143**, 29–36 (1982).

59. M. A. del Pozo, N. Balasubramanian, N. B. Alderson, W. B. Kiosses, A. Grande-García, R. G. W. Anderson, M. A. Schwartz, Phospho-caveolin-1 mediates integrin-regulated membrane domain internalization. *Nat. Cell Biol.* **7**, 901–908 (2005).

60. R. Moreno-Vicente, D. M. Pavón, I. Martín-Padura, M. Català-Montoro, A. Díez-Sánchez, A. Quílez-Álvarez, J. A. López, M. Sánchez-Álvarez, J. Vázquez, R. Strippoli, M. A. del Pozo, Caveolin-1 Modulates Mechanotransduction Responses to Substrate Stiffness through Actin-Dependent Control of YAP. *Cell Rep.* **25**, 1622-1635.e6 (2018).

61. M. Raudenska, J. Gumulec, J. Balvan, M. Masarik, Caveolin-1 in oncogenic metabolic symbiosis. *Int. J. Cancer*. **147**, 1793–1807 (2020).

62. A. Pol, F. Morales-Paytuví, M. Bosch, R. G. Parton, Non-caveolar caveolins - Duties outside the caves. *J. Cell Sci.* **133** (2020), doi:10.1242/jcs.241562.

63. H.-H. Lin, H.-K. Lin, I.-H. Lin, Y.-W. Chiou, H.-W. Chen, C.-Y. Liu, H. I.-C. Harn, W.-T. Chiu, Y.-K. Wang, M.-R. Shen, M.-J. Tang, Mechanical phenotype of cancer cells: cell softening and loss of stiffness sensing. *Oncotarget*. **6**, 20946 (2015).

64. O. Muriel, A. Echarri, C. Hellriegel, D. M. Pavón, L. Beccari, M. A. del Pozo, Phosphorylated filamin A regulates actin-linked caveolae dynamics. *J. Cell Sci.* **124**, 2763–2776 (2011).

65. R. J. Sun, S. Muller, F. Y. Zhuang, J. F. Stoltz, X. Wang, Caveolin-1 redistribution in human endothelial cells induced by laminar flow and cytokine. *Biorheology*. **40**, 31–39 (2002).

66. C. K. Hsu, H. H. Lin, H. I. Harn, R. Ogawa, Y. K. Wang, Y. T. Ho, W. R. Chen, Y. C. Lee, J. Y. Y. Lee, S. J. Shieh, C. M. Cheng, J. A. McGrath, M. J. Tang, Caveolin-1 Controls Hyperresponsiveness to Mechanical Stimuli and Fibrogenesis-Associated RUNX2 Activation in Keloid Fibroblasts. *J. Invest. Dermatol.* **138**, 208–218 (2018).

67. E. Le Master, A. Paul, D. Lazarko, V. Aguilar, S. J. Ahn, J. C. Lee, R. D. Minshall, I. Levitan, Caveolin-1 is a primary determinant of endothelial stiffening associated with dyslipidemia, disturbed flow, and ageing. *Sci. Rep.* **12**, 1–15 (2022).

68. F.-N. Lolo, N. Walani, E. Seemann, D. Zalvidea, D. M. Pavon, G. Cojoc, M. Zamai, C. Viaris de Lesegno, F. Martinez de Benito, M. Sanchez-Alvarez, J. J. Uriarte, A. Echarri, D. Jimenez-Carretero, J.-C. Escolano, S. Sanchez, V. R. Cailofa, D. Navajas, X. Trepat, J. Guck, C. Lamaze, P. Roca-Cusachs, M. Kessels, B. Qualmann, M. Arroyo, M. Á. del Pozo, Novel Caveolin1-dolines constitute a distinct and rapid mechanoadaptation system. *submitted*.

69. H.-H. Lin, H.-K. Lin, I.-H. Lin, Y.-W. Chiou, H.-W. Chen, C.-Y. Liu, H. I.-C. Harn, W.-T. Chiu, Y.-K. Wang, M.-R. Shen, M.-J. Tang, Mechanical phenotype of cancer cells: cell softening and loss of stiffness sensing. *Oncotarget*. **6** (2015), doi:10.18632/oncotarget.4173.

70. H. A. Hirsch, D. Iliopoulos, P. N. Tsichlis, K. Struhl, Metformin selectively targets cancer stem cells, and acts together with chemotherapy to block tumor growth and prolong remission. *Cancer Res.* **69**, 7507–7511 (2009).

71. S. Tavares, A. F. Vieira, A. V. Taubenberger, M. Araújo, N. P. Martins, C. Brás-Pereira, A. Polónia, M. Herbig, C. Barreto, O. Otto, J. Cardoso, J. B. Pereira-Leal, J. Guck, J. Paredes, F. Janody, Actin

stress fiber organization promotes cell stiffening and proliferation of pre-invasive breast cancer cells. *Nat. Commun.* **8**, 1–18 (2017).

72. H. A. Hirsch, D. Iliopoulos, A. Joshi, Y. Zhang, S. A. Jaeger, M. Bulyk, P. N. Tsichlis, X. Shirley Liu, K. Struhl, A Transcriptional Signature and Common Gene Networks Link Cancer with Lipid Metabolism and Diverse Human Diseases. *Cancer Cell*. **17**, 348–361 (2010).

73. M. Urbanska, H. E. Muñoz, J. Shaw Bagnall, O. Otto, S. R. Manalis, D. Di Carlo, J. Guck, A comparison of microfluidic methods for high-throughput cell deformability measurements. *Nat. Methods*. **17**, 587–593 (2020).

74. C. Durán, S. Ciucci, A. Palladini, U. Z. Ijaz, A. G. Zippo, F. P. Sterbini, L. Masucci, G. Cammarota, G. Ianiro, P. Spuul, M. Schroeder, S. W. Grill, B. N. Parsons, D. M. Pritchard, B. Posteraro, M. Sanguinetti, G. Gasbarrini, A. Gasbarrini, C. V. Cannistraci, Nonlinear machine learning pattern recognition and bacteria-metabolite multilayer network analysis of perturbed gastric microbiome. *Nat. Commun.* **12**, 1926 (2021).

75. R. Strippoli, P. Sandoval, R. Moreno-Vicente, L. Rossi, C. Battistelli, M. Terri, L. Pascual-Antón, M. Loureiro, F. Matteini, E. Calvo, J. A. Jiménez-Heffernan, M. J. Gómez, V. Jiménez-Jiménez, F. Sánchez-Cabo, J. Vázquez, M. Tripodi, M. López-Cabrera, M. Á. del Pozo, Caveolin1 and YAP drive mechanically induced mesothelial to mesenchymal transition and fibrosis. *Cell Death Dis.* **11**, 1–19 (2020).

76. S. Dupont, L. Morsut, M. Aragona, E. Enzo, S. Giulitti, M. Cordenonsi, F. Zanconato, J. Le Digabel, M. Forcato, S. Bicciato, N. Elvassore, S. Piccolo, Role of YAP/TAZ in mechanotransduction. *Nature*. **474**, 179–184 (2011).

77. V. Rausch, J. R. Bostrom, J. Park, I. R. Bravo, Y. Feng, D. C. Hay, B. A. Link, C. G. Hansen, The Hippo Pathway Regulates Caveolae Expression and Mediates Flow Response via Caveolae. *Curr. Biol.* **29**, 242-255.e6 (2019).

78. C. Stein, A. F. Bardet, G. Roma, S. Bergling, I. Clay, A. Ruchti, C. Agarinis, T. Schmelzle, T. Bouwmeester, D. Schübeler, A. Bauer, YAP1 Exerts Its Transcriptional Control via TEAD-Mediated Activation of Enhancers. *PLoS Genet.* **11**, 1–28 (2015).

79. B. Zhao, X. Ye, J. Yu, L. Li, W. Li, S. Li, J. Yu, J. D. Lin, C. Y. Wang, A. M. Chinnaiyan, Z. C. Lai, K. L. Guan, TEAD mediates YAP-dependent gene induction and growth control. *Genes Dev.* **22**, 1962–1971 (2008).

80. W. R. Jiang, G. Cady, M. M. Hossain, Q. Q. Huang, X. Wang, J. P. Jin, Mechanoregulation of h2-calponin gene expression and the role of notch signaling. *J. Biol. Chem.* **289**, 1617–1628 (2014).

81. M. M. Hossain, J. F. Crish, R. L. Eckert, J. J. C. Lin, J. P. Jin, H2-Calponin Is Regulated By Mechanical Tension and Modifies the Function of Actin Cytoskeleton. *J. Biol. Chem.* **280**, 42442–42453 (2005).

82. M. De Marzio, A. Kiliç, E. Maiorino, J. A. Mitchel, C. Mwase, M. J. O'Sullivan, M. McGill, R. Chase, J. J. Fredberg, J. A. Park, K. Glass, S. T. Weiss, Genomic signatures of the unjamming transition in compressed human bronchial epithelial cells. *Sci. Adv.* **7**, 1–16 (2021).

83. A. Kılıç, A. Ameli, J. A. Park, A. T. Kho, K. Tantisira, M. Santolini, F. Cheng, J. A. Mitchel, M. McGill, M. J. O'Sullivan, M. De Marzio, A. Sharma, S. H. Randell, J. M. Drazen, J. J. Fredberg, S. T. Weiss, Mechanical forces induce an asthma gene signature in healthy airway epithelial cells. *Sci. Rep.*

10, 1–12 (2020).

84. J. Rysä, H. Tokola, H. Ruskoaho, Mechanical stretch induced transcriptomic profiles in cardiac myocytes. *Sci. Rep.* **8**, 1–14 (2018).

85. J. Lv, Y. Liu, F. Cheng, J. Li, Y. Zhou, T. Zhang, N. Zhou, C. Li, Z. Wang, L. Ma, M. Liu, Q. Zhu, X. Liu, K. Tang, J. Ma, H. Zhang, J. Xie, Y. Fang, H. Zhang, N. Wang, Y. Liu, Cell softness regulates tumorigenicity and stemness of cancer cells, 1–14 (2020).

86. J. G. Goetz, P. Lajoie, S. M. Wiseman, I. R. Nabi, Caveolin-1 in tumor progression: The good, the bad and the ugly. *Cancer Metastasis Rev.* **27**, 715–735 (2008).

87. Z. Wang, N. Wang, P. Liu, F. Peng, H. Tang, Q. Chen, R. Xu, Y. Dai, Y. Lin, X. Xie, C. Peng, H. Situ, Caveolin-1, a stress-related oncotarget, in drug resistance. *Oncotarget*. **6**, 37135–37150 (2015).

88. A. A. A. A. Nawaz, M. Urbanska, M. Herbig, M. Nötzel, M. Kräter, P. Rosendahl, C. Herold, N. Toepfner, M. Kubánková, R. Goswami, S. Abuhattum, F. Reichel, P. Müller, A. Taubenberger, S. Girardo, A. Jacobi, J. Guck, Intelligent image-based deformation-assisted cell sorting with molecular specificity. *Nat. Methods*. **17**, 595–599 (2020).

89. A. A. Nawaz, D. Soteriou, C. K. Xu, R. Goswami, M. Herbig, J. Guck, S. Girardo, Image-based cell sorting using focused travelling surface acoustic waves. *Lab Chip*, 372–387 (2023).

90. B. Razani, J. A. Engelman, X. B. Wang, W. Schubert, X. L. Zhang, C. B. Marks, F. Macalusol, R. G. Russell, M. Li, R. G. Pestell, D. Di Vizio, H. Hou, B. Kneitz, G. Lagaud, G. J. Christ, W. Edelmann, M. P. Lisanti, Caveolin-1 Null Mice Are Viable but Show Evidence of Hyperproliferative and Vascular Abnormalities. *J. Biol. Chem.* **276**, 38121–38138 (2001).

91. M. Urbanska, P. Rosendahl, M. Kräter, J. Guck, High-throughput single-cell mechanical phenotyping with real-time deformability cytometry. *Methods Cell Biol.* **147** (2018), doi:10.1016/bs.mcb.2018.06.009.

92. I. N. Sneddon, The relation between load and penetration in the axisymmetric boussinesq problem for a punch of arbitrary profile. *Int. J. Eng. Sci.* **3**, 47–57 (1965).

93. J. Alcaraz, L. Buscemi, M. Grabulosa, X. Trepat, B. Fabry, R. Farré, D. Navajas, Microrheology of human lung epithelial cells measured by atomic force microscopy. *Biophys. J.* **84**, 2071–2079 (2003).

94. J. Rother, H. Nöding, I. Mey, A. Janshoff, Atomic force microscopy-based microrheology reveals significant differences in the viscoelastic response between malign and benign cell lines. *Open Biol.* (2014), doi:10.1098/rsob.140046.

95. J. Alcaraz, L. Buscemi, M. Puig-De-Morales, J. Colchero, A. Baró, D. Navajas, Correction of microrheological measurements of soft samples with atomic force microscopy for the hydrodynamic drag on the cantilever. *Langmuir* (2002), doi:10.1021/la0110850.

96. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape : A Software Environment for Integrated Models of Biomolecular Interaction Networks, 2498–2504 (2003).

97. P. Danaher, S. Warren, L. Dennis, L. D'Amico, A. White, M. L. Disis, M. A. Geller, K. Odunsi, J. Beechem, S. P. Fling, Gene expression markers of Tumor Infiltrating Leukocytes. *J. Immunother. Cancer*. **5**, 1–15 (2017).

98. A. Acevedo, C. Duran, M. J. Kuo, S. Ciucci, M. Schroeder, C. V. Cannistraci, Measuring Group Separability in Geometrical Space for Evaluation of Pattern Recognition and Dimension Reduction Algorithms. *IEEE Access*. **10**, 22441–22471 (2022).

99. M. Herbig, A. Mietke, P. Müller, O. Otto, Statistics for real-time deformability cytometry: Clustering, dimensionality reduction, and significance testing. *Biomicrofluidics*. **12** (2018), doi:10.1063/1.5027197.

100. C. Herold, http://arxiv.org/abs/1704.00572 (2017).

101. M. Itoh, M. Kojima, S. Nagao-Sato, E. Saijo, T. Lassmann, M. Kanamori-Katayama, A. Kaiho, M. Lizio, H. Kawaji, P. Carninci, A. R. R. Forrest, Y. Hayashizaki, Automated workflow for preparation of cDNA for cap analysis of gene expression on a single molecule sequencer. *PLoS One*. **7** (2012), doi:10.1371/journal.pone.0030809.