Volumetric compression develops noise-driven single-cell heterogeneity

Xing Zhao[a,b,c,1], Jiliang Hu[d,1], Yiwei Li[a,2], and Ming Guo[d,2]

[a]Department of Biomedical Engineering, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China; [b]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138; [c]BGI-Shenzhen, Shenzhen 518083, China; and [d]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by Herbert Levine, Northeastern University, Boston, MA; received June 7 2021; accepted November 2, 2021

Recent studies have revealed that extensive heterogeneity of biological systems arises through various routes ranging from intracellular chromosome segregation to spatiotemporally varying biochemical stimulations. However, the contribution of physical microenvironments to single-cell heterogeneity remains largely unexplored. Here, we show that a homogeneous population of non–small-cell lung carcinoma develops into heterogeneous subpopulations upon application of a homogeneous physical compression, as shown by single-cell transcriptome profiling. The generated subpopulations stochastically gain the signature genes associated with epithelial–mesenchymal transition (EMT; VIM, CDH1, EPCAM, ZEB1, and ZEB2) and cancer stem cells (MKI67, BIRC5, and KLF4), respectively. Trajectory analysis revealed two bifurcated paths as cells evolving upon the physical compression, along each path the corresponding signature genes (epithelial or mesenchymal) gradually increase. Furthermore, we show that compression increases gene expression noise, which interplays with regulatory network architecture and thus generates differential cell-fate outcomes. The experimental observations of both single-cell sequencing and single-molecule fluorescent in situ hybridization agrees well with our computational modeling of regulatory network in the EMT process. These results demonstrate a paradigm of how mechanical stimulations impact cell-fate determination by altering transcription dynamics; moreover, we show a distinct path that the ecology and evolution of cancer interplay with their physical microenvironments from the view of mechanobiology and systems biology, with insight into the origin of single-cell heterogeneity.

cell volume | mechanobiology | single cell | heterogeneity | cell fate decision

A multicellular biological system, such as organ, tumor, and embryo, is composed of a complex network of heterogeneous yet closely related cells (1). Biological heterogeneity is well organized on several orders of scales and is always originated from a single homogeneous population of cells (2–4). Recent studies have shown that extensive genetic and phenotypic variations exist not only between cell types but also within a seemingly homogenous population, which suggest a hint of the origin and causality of cellular heterogeneity. Indeed, various biological and genetic origins of heterogeneity have been revealed, including genomic instability (5, 6), stochastic gene expression (7–9), heterotypic chromosome segregation (10), and epithelial–mesenchymal transition (EMT) (11) as well as spatiotemporally regulated stimulations of morphogens (12). Despite extensive studies of biological and genetic causes of genetic and phenotypic heterogeneity, little is known about the impact of physical microenvironments on heterogeneous cell-fate decision (13, 14). Furthermore, the hallmarks of cancer include specific changes of mechanical properties of cancer cells and their physical microenvironments such as increased extracellular matrix stiffness (15–22), elevated osmotic stresses (23–27), and mechanical forces (28–31). All these physical parameters have been shown to generate a wide spectrum of downstream effects including activation of tumorigenic signaling [e.g., YAP/TAZ (32, 33) and β-catenin (28)], promotion of tumor invasion (15, 17, 27, 29, 31, 34), and hyperproliferation of cancer cells (28). Nevertheless, whether these physical perturbations can directly impact the heterogeneous nature on the single-cell level remains largely unknown (35, 36).

Toward this end, we directly track single-cell gene expression within a homogeneous population of non–small-cell lung carcinoma using high-throughput single-cell sequencing, as we apply a homogeneous physical compression. We clearly observe the generation of subpopulations with a pattern that several clusters of cells separate from each other by visualizing single-cell transcriptome. The genes associated with EMT (VIM, CDH1, EPCAM, ZEB1, and ZEB2) (37–40) and cancer stem cells (MKI67, BIRC5, and KLF4) (41–46) begin to gradually self-correlate among the newly generated subpopulations. A trajectory analysis using partition-based graph abstraction (PAGA) (47) shows two distinct evolution paths of non–small-cell lung carcinoma upon the compression. The two paths are termed "mesenchymal path" and "epithelial path," respectively, as their corresponding genes (mesenchymal or epithelial) gradually increase along the pseudotime. This unique behavior of two cell-fate outcomes with one environmental input is well explained by the dynamical bifurcation in our EMT regulatory model with multistability. The experimental

Significance

Tumor heterogeneity is widely attributed to the imperfection of DNA replication. However, little is known about the mechanoregulation of tumor heterogeneity. Here, we report that volumetric compression that usually arises from tumor progression increases the overall gene-expression noise, leading to differential cell-fate transitions along epithelial/mesenchymal transition regulatory network from homogeneous non–small-cell lung carcinoma. The increased noise could be caused by a transit decrease in gene expression following the decreasing cell volume under compression. Both the experiments and numerical modeling confirmed the differential cell-fate transitions from the hybrid epithelial/mesenchymal state to either epithelial or mesenchymal states stochastically. Thus, we suggest that the cause of tumor heterogeneity could be its mechanical microenvironment as sensed by its cytoplasmic volume.

Author contributions: Y.L. and M.G. designed research; X.Z. and Y.L. performed research; X.Z., J.H., and Y.L. analyzed data; X.Z., J.H., Y.L., and M.G. wrote the paper; and J.H. conducted numerical modeling and simulation.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]X.Z. and J.H. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: guom@mit.edu or yiweili@hust.edu.cn.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2110550118/-/DCSupplemental.

Published December 16, 2021.

validations using single-cell sequencing data and single-molecule RNA fluorescence in situ hybridization (sm-FISH) confirm that the gene expression noise is elevated after compression and leads to generation of new heterogeneous subgroups according to EMT regulatory network.

Results

Physical Compression Develops Single-Cell Heterogeneity in a Homogeneous Cell Population. To explore the impact of physical compression on single-cell heterogeneity development, we employ the current high-throughput single-cell sequencing technology (48–51) to track the single-cell transcriptome evolution among a homogeneous population of non–small-cell lung carcinoma (Fig. 1*A*). Lung carcinoma is a malignant lung tumor characterized by uncontrolled cell growth in lung tissues (52). In lung tissues, the dysregulated osmotic pressure is associated with many lung diseases, including cystic fibrosis, pleural effusion, and cancers (24, 53–55). To investigate the effect of elevated osmolarity on cell-fate decision in developing lung tumors, we treat non–small-cell lung carcinoma H1975 cells with hypertonic medium containing 2% PEG 300 (Polyethylene glycol, 300 MW), which is an inert polymer that is widely used to avoid nonspecific interactions with biomolecules (56, 57). At different time points (2.5 h or 4 h) after the treatment of osmotic compression, cells are harvested for single-cell sequencing analysis using commercial 10× chromium platform (49) (Fig. 1*A*). To classify major cell subpopulation generation, we perform uniform manifold approximation and projection (UMAP) (58) analysis using Scanpy (59) and have identified two major groups arising upon osmotic compression (Fig. 1*B*). To further analyze the subclusters within each major cell population, we segregate cells into several different subtypes according to differentially expressed genes. By analyzing differentially expressed genes among the subclusters, we identify several signature genes associated with EMT process (VIM, CDH1, EPCAM, CTGF, and LOX) (37, 40) and cancer stem cells (MKI67, BIRC5, and KLF4) (41–44) (Figs. 1*C* and 2 and *SI Appendix*, Fig. S1). To confirm that the EMT-related subpopulations are indeed generated upon osmotic compression, we compare the Pearson correlation coefficients for both the total genes (Fig. 1*D*) and the gene set of EMT process before and after compression, including 14 signature genes representing epithelial phenotype and 7 genes representing mesenchymal phenotype (Fig. 1*E*). In general, we observe strong positive correlation coefficients among their own populations after compression (Fig. 1 *D* and *E*).

Furthermore, to examine that the newly generated subpopulations are not a transient state occurring on the transcription level and last to form new phenotypes on the protein expression level, we apply fluorescent immunostaining assay to visualize the expression level of epithelial marker E-cadherin and mesenchymal marker Vimentin (60) at day 0 and day 2 upon osmotic compression (~400 mOsm, 2% PEG). As shown in Fig. 1*F*, before osmotic compression, H1975 cells express comparable amount of both E-cadherin and Vimentin. After 2 d of stimulation of hypertonic pressure, cells are prone to express either E-cadherin or Vimentin but not both (Fig. 1*F*). The outcomes of differential expressions of either E-cadherin or Vimentin in individual cells are consistent with our single cell transcriptome analysis (Fig. 1 *C* and *F*). In addition, we also notice that the phenotypic changes on the protein expression level are longer than that which we observe on the messenger RNA (mRNA) level.

Reconstruction of Hierarchy of Compression-Induced Heterogeneity. Our experimental studies focus on cells at different time points after the osmotic compression. These newly generated subpopulations and the transitions between them can be computationally reconstructed using trajectory analysis. We perform PAGA (47), which reconciles clustering and pseudotemporal ordering algorithms and allows the inference of complex cell trajectories and differentiation trees (61–63). Using this algorithm, we obtain an abstracted graph that shows high confidence of the branching events (Fig. 2*A*), from which we derive a single differentiation tree that includes all the cell types and links them to a single root—the cluster of cells without compression. This tree defines independent, bifurcated differentiation branches of mesenchymal path and epithelial path (Fig. 2*B*). After 2.5 h of osmotic compression, the cells from the same group begin to separately locate on the two sides of the root cells (without compression); the same trend has also been confirmed in the cells after 4 h of compression (Fig. 2 *A* and *B*). The connections in the tree are highly consistent with the continuity of gene-expression patterns along the two different paths. Furthermore, PAGA yields a pseudotemporal ordering of individual cells within each path, consistent with our cells cultured after different durations of compression (Fig. 2*B*). This pseudotemporal ordering reveals a gradual increase in mesenchymal genes in the mesenchymal path; these mesenchymal genes gradually decrease in the epithelial path (Fig. 2 *C–E* and *I*). Meanwhile, the epithelial genes gradually decrease along the mesenchymal path but increase along the epithelial path (Fig. 2 *F–I*). In addition to the PAGA algorithm, we also cluster and order the single-cell sequencing data using other embedding methods including UMAP, t-distributed stochastic neighbor embedding (t-SNE), and diffusion map, which consistently support our conclusion generated from PAGA ordering (*SI Appendix*, Figs. S2–S4). Furthermore, we have also tested markers for the hybrid E/M cells, such as NRF2 (64). As shown in *SI Appendix*, Fig. S5, we clearly observe that NRF2 is maximally expressed in the initial control group without compression. These results confirm that the initial control group is indeed the hybrid E/M cells.

Since the generation of new subgroups is found to be related to the EMT phenotypes, we further test whether the EMT signature genes can efficiently recapitulate the newly generated subgroups. The list of EMT signature genes is obtained from previous studies (39, 40, 65). After filtering, 37 genes are selected and sufficiently recapitulate the newly generated subgroups (*SI Appendix*, Figs. S6 and S7). Interestingly, as we further narrow down the gene list, we find that even five genes or three genes are sufficient to cluster cell subgroups (*SI Appendix*, Figs. S8 and S9). The three genes used are VIM, CDH1, and CLDN7, which may be the minimal set of genes for separating EMT subgroups. This result is consistent with the multinomial logistic regression–based EMT scoring method (40, 65), which reported that the relationship between VIM/CDH1 ratio and CLDN7 was able to distinguish different subgroups in EMT phenotypes.

Compression Alters Transcription Rates and Increases Gene-Expression Noise during Development of Heterogeneity Hierarchy. The interesting consequence of the compression as shown above is the promotion of two opposite phenotypes that suppress each other. Biochemical stimulation usually generates a deterministic effect: For example, TGF-β is well known for induction of EMT process (66, 67); meanwhile, mesenchymal-to-epithelial transition, the reverse process of EMT, can be induced by the stimulation of Bone morphogenic protein-7 (BMP-7) (68). The dual-elevation effect of volumetric compression on two opposite processes suggests a distinct underlying mechanism that differs from biochemical regulations. Instead, we speculate that the volumetric compression results in global changes in overall transcription kinetics, which eventually lead to differential cell-fate decisions. In addition, previous studies have also shown a similar behavior in

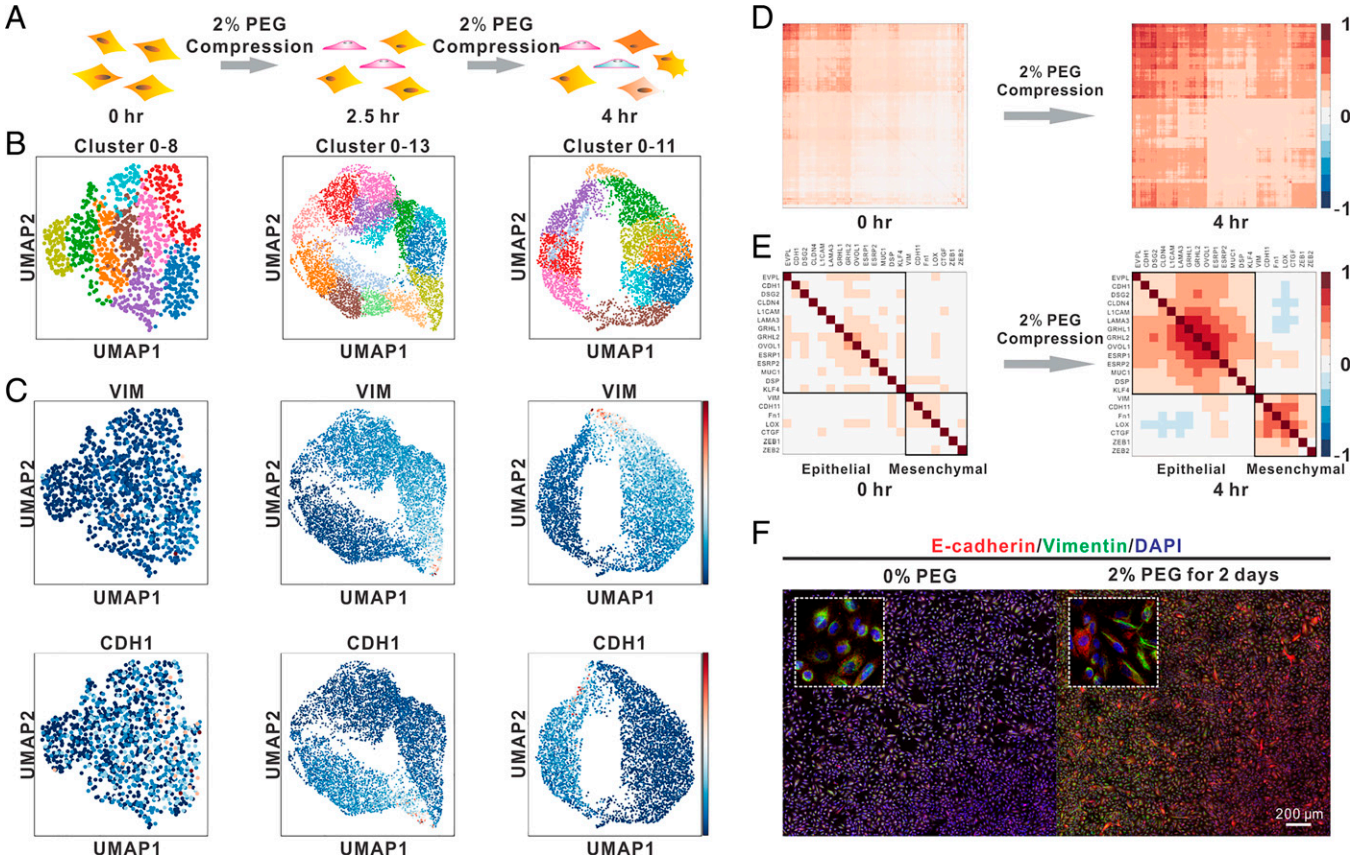

Fig. 1. Physical compression develops heterogeneous subpopulations with signatures of epithelial genes and mesenchymal genes, respectively, from a seemly homogeneous population of non–small-cell lung carcinoma. (*A*) Schematic illustration of application of hypertonic medium on non–small-cell lung carcinoma H1975. (*B*) UMAP plot of non–small-cell lung carcinoma H1975 cells under different time periods of osmotic compression to visualize heterogeneous subpopulations development. (*C*) Mesenchymal marker VIM (*Top*) and Epithelial marker CDH1 (*Bottom*). The relative expression level of genes across cells is shown in UMAP plot. (*D*) Correlation of all detected genes expression between individual cells before and after 4 h of compression. The result indicates a stronger correlation strength after 4 h of compression. (*E*) Correlation of epithelial genes and mesenchymal genes expression between individual cells before and after 4 h of compression. The result indicates a stronger correlation strength within the group of epithelial genes and mesenchymal gens respectively after 4 h of compression. (*F*) Immune costaining of epithelial protein marker E-cadherin and mesenchymal protein marker Vimentin before and after 2 d of compression. The cells before osmotic compression expressing comparable and mild amounts of E-cadherin and Vimentin. After 2 d of compression, the cells prone to express larger amount of one kind of marker. (Scale bar, 200 μm.)

prokaryote (69–73) and some mammalian cells (8, 74–81), in which the gene-expression noise can drive a bistable switch of cell-fate states in a mutual antagonism regulation network topology (82–84). Furthermore, it has also been recently reported that variance (noise) can serve as an early warning signal for the critical transition in mammalian cells, specifically in the EMT regulatory network (85). Nevertheless, a key question in differential cell-fate transitions of mammalian cells is the following: What cue or regulator in cell microenvironments is responsible for the gene-expression noise modulation and thus switching on the multistate cell-fate transition (86–90).

As inspired by the above-mentioned studies, we further explore whether the changes of gene-expression abundance and noise (variation) can be found in cells upon volumetric compression. Previous studies have revealed that osmotic compression tightly regulated the cellular volume (91–93). Another study demonstrated a linear correlation between cell volumes and mRNA counts to maintain a consistent concentration of mRNA (76, 79, 94). Thus, we test whether the volumetric compression by osmotic stresses also alters intrinsic transcription kinetics and changes transcript abundance. To do so, we first conduct sm-FISH to quantitatively count the numbers of transcripts of selected ubiquitously expressed genes, which more closely follow the volume-dependent transcription and better reflect intrinsic transcription kinetics. Genes, including VIM, CDH1, and GAPDH, are firstly tested (Fig. 3 *A*–*D*). This assay is performed at different time points (0 min, 15 min, 30 min, 45 min, 1 h, 2 h, 3 h, and 4 h) after the application of volumetric compression. Accompanying with the decreased cell volume, we first observe a decreased number of mRNA transcripts upon osmotic compression from the beginning 30 mins (Fig. 3 *B*–*D*), which is consistent with volume-dependent expression reported previously (76, 79). However, the VIM and CDH1 mRNA transcripts start to increase 1 h after the application of volumetric compression (Fig. 3 *B* and *C*). The number of VIM mRNA can reach a value that is similar to the number of transcripts before compression, while the number of CDH1 mRNA increases 30% after 4 h of osmotic compression. Furthermore, we directly compare the gene noises of tested genes before and after volumetric compression by calculating the coefficient of variation of the mRNA transcript number (95). Practically, four EMT-related genes (VIM, ZEB1, ZEB2, and CDH1) (37) are selected for the following sm-FISH assay. After 30 mins of compression, we observe an increase in gene-expression noise, which is consistent with the above-mentioned time-dependent measurements (Fig. 3 *E*–*H*). This result confirms that the volumetric compression, at least transiently, decreases the expression of EMT-related genes. This decreased gene expression

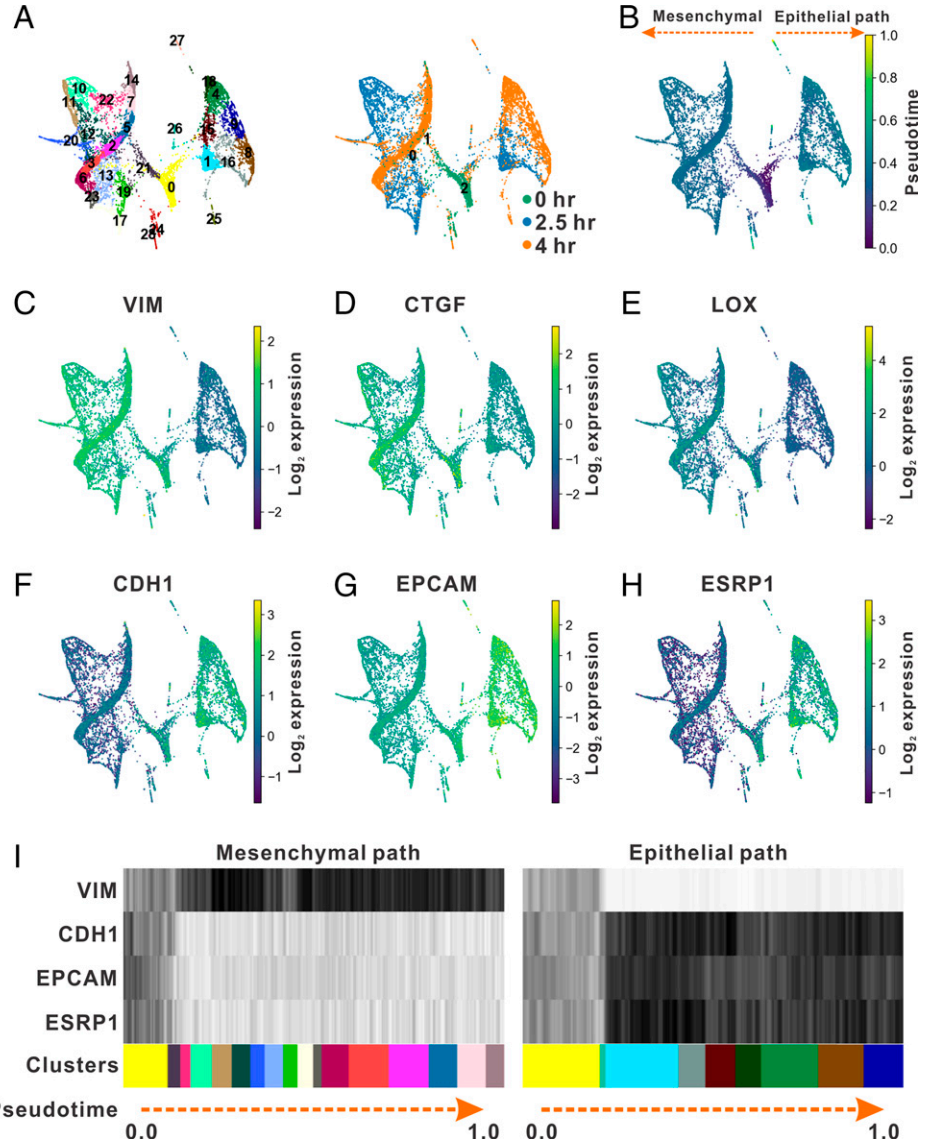

Fig. 2. Trajectory analysis reconstructs the hierarchy of compression-induced heterogeneity. (*A*) Mapping the topology of non–small-cell lung carcinoma. The cells were clustering using the Louvain algorithm, and the relationship of clusters was measured using the PAGA algorithm. (*Right*) Cells categorized by their time periods of compression treatment are shown in PAGA plot. (*B*) The sequential color indicates that the position of each cell in pseudotime starts from 0 to 1 in PAGA layout. (*C–E*) The expression of mesenchymal genes in each cell is shown in PAGA layout. (*F–H*) The mean expression of epithelial genes in each cell is shown in PAGA layout. (*I*) Expression kinetics of selected maker genes along pseudotime.

also accompanies with the increased gene-expression noise, which is believed to follow the Poisson noise. For a longer period of time (4 h) of volumetric compression, the average number of mRNA transcripts of EMT-related genes turns out to be increased, following a dramatic increase in the gene-expression noise (Fig. 3 *B* and *C*). To confirm the increased noise is also observed in our single-cell sequencing data, we calculate the overall coefficient of variation of all detected genes, which consistently shows an increased gene-expression noise after volumetric compression (Fig. 3*I*). In addition to the increased variation observed in the mRNA level, we also calculate the uncorrelated noise from the proteins (Vimentin and E-cadherin) immunostaining result and the variation in different cell morphological parameters (*SI Appendix*, Figs. S10–S12) (35, 84, 96–100). We find that both the uncorrelated noise in protein expression and the variation in 13 morphological parameters are increased upon compression, which support our conclusion that volumetric compression increases the variation (noise) and subpopulation generation in H1975 cells.

Based on the obtained sm-FISH data at different time points, we first confirm that the volumetric compression increases gene-expression noise. This compression-induced variation on the expression of EMT-related genes latterly leads to a differential distribution of the numbers of EMT-related mRNAs, giving rise to subpopulation generation. These observed sequential events of gene noise increasing and subpopulation dividing together suggest an important role of the interplay between gene-expression noise and regulatory network topology in responding to physical compression.

The Multistability of Cell Fate Emerges from the Dynamical Bifurcation of EMT Regulatory Network. To further understand how the increased noise leads to differential cell-fate outcomes, we construct a core EMT regulatory network based on the full EMT

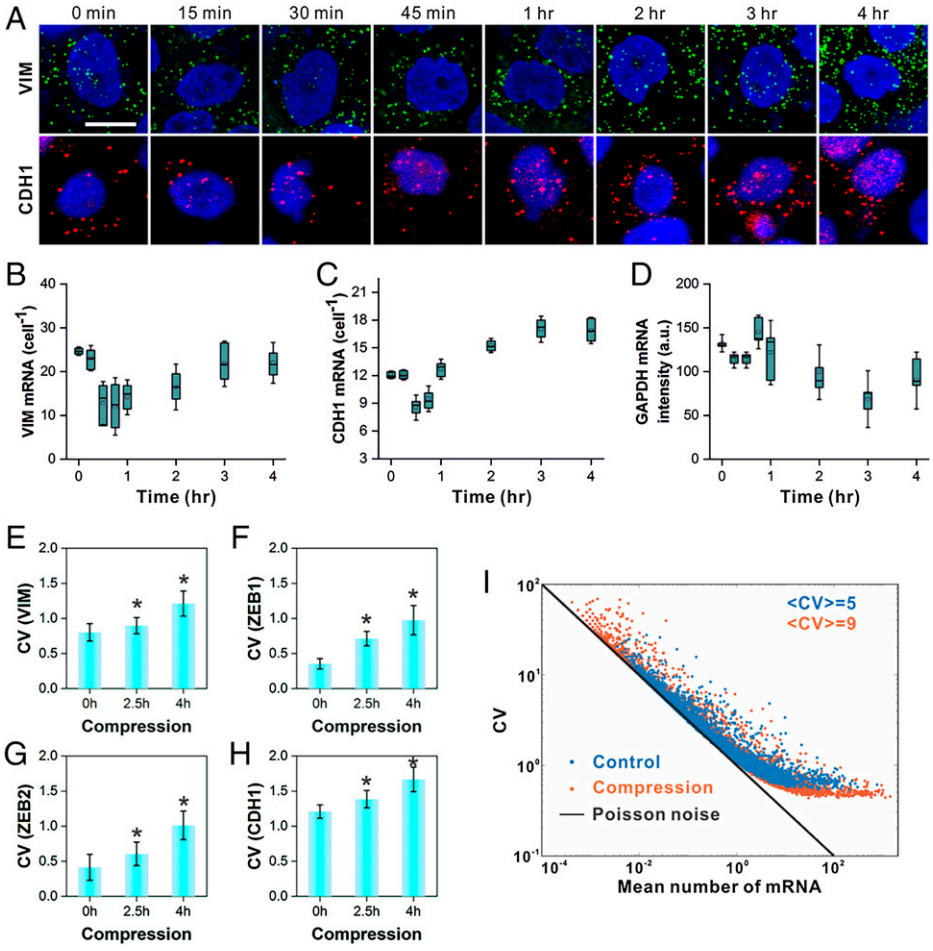

Fig. 3. Compression increases the gene-expression noise indicated by sm-FISH. (*A*) Images of sm-FISH of VIM and CDH1 at different time points post volumetric compression. (Scale bar, 5 μm.) (*B* and *C*) Quantification of VIM (*B*) and CDH1 (*C*) mRNA counts at different time points post volumetric compression. $n = 5$ independent experiments, and each independent experiment of mRNA measurement contains 15 cells. (*D*) Quantification of GAPDH mRNA expression at different time points post volumetric compression. (*E*–*H*) Quantification of gene-expression noises of epithelial/mesenchymal marker genes, indicating increased gene-expression noise after compression. The marker genes include VIM (*E*), ZEB1 (*F*), ZEB2 (*G*), and CDH1 (*H*). $n = 240$ cells from three independent experiments in each condition. (*I*) The overall gene-expression noise was increased after compression as indicated from single-cell sequencing data.

regulatory network reported in previous literatures (18, 67, 101–105). Without loss of generality, we consider the antagonistic interaction between a collection of miR-34 and miR-200 and a collection of SNAIL and ZEB1; the collection of SNAIL and ZEB1 directly inhibits CDH1, while promoting the expression of VIM (103), as illustrated in Fig. 4*A*. A more accurate EMT regulatory network will incorporate many intermediate genes transmitting the interaction of the genes in our model, which still yields the same qualitative results that we show here (102, 103). The dynamics of the core EMT network can be calculated using equations considering two mutually opposing transcription factors (TFs) with positive self-regulations (106–108). For our purpose, it is sufficient to consider a symmetric parameter condition here, which means $a_1 = a_2$, $b_1 = b_2$, $k_1 = k_2$.

$$\begin{aligned}\frac{dx_1}{dt} &= \frac{a_1 x_1^n}{S^n + x_1^n} + \frac{b_1 S^n}{S^n + x_2^n} - k_1 x_1 = F_1(x_1, x_2)\\ \frac{dx_2}{dt} &= \frac{a_2 x_2^n}{S^n + x_2^n} + \frac{b_2 S^n}{S^n + x_1^n} - k_2 x_2 = F_2(x_1, x_2)\end{aligned}, \qquad [1]$$

where x_1 and x_2 are the concentrations of TF 1 (a collection of miR-34 and miR-200) and TF 2 (a collection of SNAIL1 and ZEB1), respectively. On the right side of the equations, the first terms represent the positive self-regulation with strength a_1 and a_2, the second terms represent mutual inhibition with a basal expression strength of b_1 and b_2, the third terms represent the degradation of each TF with rate k_1 and k_2, the exponent n, known as the Hill coefficient, is a measure of the cooperativity of the response, and S is the dissociation constant measuring the steepness of the first term sigmoidal function of self-regulation. Stationary solutions of x_1 and x_2 occur when dx_1/dt and dx_2/dt both equal zero (i.e., $F_1(x_1, x_2)$ and $F_2(x_1, x_2)$ both equal zero).

The transition from the hybrid state before the application of volumetric compression to the two distinct epithelial and mesenchymal states after compression is driven by the decreasing gene-expression rate and the increasing gene-expression noise (Fig. 4 *B* and *C*). As the bifurcation diagram in Fig. 4*B* shows, the size of the central attractor basin of hybrid state gradually decreases with the decrease of self-regulation with strength a_1 and a_2; a bifurcation happens at $a=a_{\text{critical}}$ when we further decrease a_1 and a_2, which corresponds to the disappearance of the hybrid state, and each cell follows either the epithelial state or the mesenchymal state. The decrease of transcription rate of x_1 and x_2 is the driving force of the bifurcation, which is shown in Fig. 4*C*.

To incorporate the stochastic gene-expression noise and develop the cell-fate landscape, we describe the evolution of probability distribution P (x_1, x_2, t) with Fokker–Planck

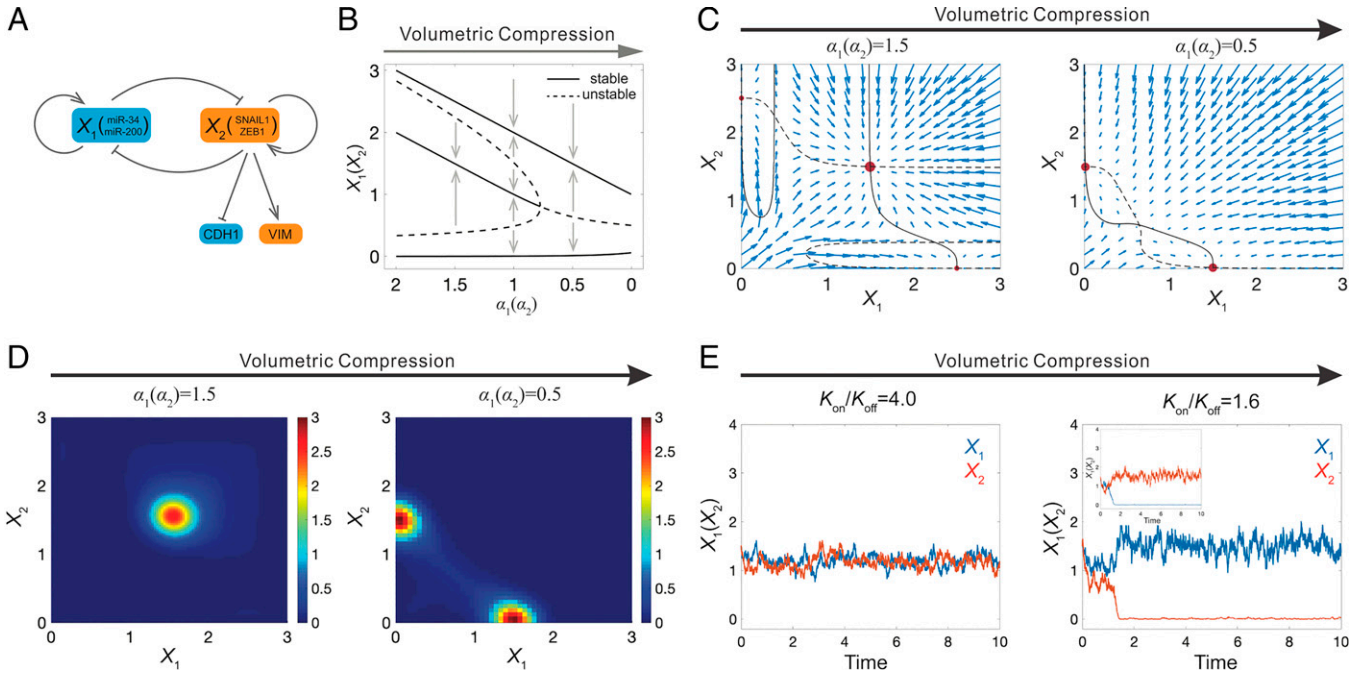

Fig. 4. Model of the bifurcation dynamics of gene expression and quantitative explanation of the transition of cell fate with decrease of transcription rate across a critical point. Bifurcation analysis, phase portrait, Fokker–Planck equation, and stochastic simulation provide evidence that compression develops single-cell heterogeneity. (*A*) Schematic representations of the mutual antagonism in EMT regulatory network. (*B*) Bifurcation diagram indicates the destabilization of central hybrid state and stabilization of epithelial and mesenchymal states when symmetrically decreasing a1 and a2. (*C*) The phase portrait shows the vector field in state space for parameters before and after the critical bifurcation point acritical. (*D*) Steady-state probability distribution calculated from Fokker–Planck equation for parameters before and after the critical bifurcation point acritical. (*E*) The evolution of TFs simulated with the Gillespie algorithm, which shows the emergent of E/M states from H state with the decrease of Kon/Koff. The insert indicates another possible consequence.

equation, in which diffusion coefficient D represents the genetic noise (109–111). On the right side of the equation, the first two terms represent the TF generation driving drift in the state space, and the last two terms represent noise driving diffusion in the state space:

$$\frac{\partial P(x_1,x_2,t)}{\partial t} = -\frac{\partial}{\partial x_1}[F_1(x_1,x_2)*P] - \frac{\partial}{\partial x_2}[F_2(x_1,x_2)*P] + \frac{\partial^2}{\partial {x_1}^2}[D(x_1,x_2)*P] + \frac{\partial^2}{\partial {x_2}^2}[D(x_1,x_2)*P] \quad [2]$$

As shown in Fig. 4*D*, as expression rate a_1 and a_2 decrease, the central hybrid state is destabilized (decreasing probability) as the marginal epithelial and mesenchymal states are gradually stabilized (increasing probability).

The transcription in cell nucleus happens as stochastic bursts (112–114); the burst size and frequency decrease with a lower TF enhancer/promotor accessibility. The volume-correlated regulation can change either the burst frequency or the burst size by modifying TF binding efficiency to enhancer/promotor. To directly study the effect of decreasing transcription burst size/frequency on bifurcated cell-fate decision, we perform stochastic simulation on the regulatory network dynamics with the Gillespie algorithm. We use K_{on} and K_{off} to describe the binding and dissociation of enhancer/promotor to TFs, respectively; the results show that the epithelial and mesenchymal states appear with the decrease of K_{on}/K_{off}, as shown in Fig. 4*E*.

We also test whether only self-activation for one of the two nodes (x_1, x_2) is enough to generate qualitative similar patterns in Fig. 4, as considering the mesenchymal–epithelial transition (MET) inducer could be weaker than that of EMT process. Here, we replace the nonlinear self-activation term for x_1 (first term containing a_1 in the Eq. **1**) with a constant term of c_1, which means that x_1 only exhibits constant expression rate without any self-activation. The rest of the equations remain the same; thus, we only consider the self-activation of EMT inducers here. We observe that the hybrid state still dominates the cell fate under large gene-expression rate regime ($a_2 = 1.5$ and $c_1 = 1.5$) as shown in the left panels of *SI Appendix*, Fig. S13. As we decrease gene-expression rate ($a_2 = 0.5$ and $c_1 = 0.3$) as applying compression, the hybrid state disappears, and the cell fates eventually fall into either x_1 dominating or x_2 dominating state (corresponding to epithelial and mesenchymal states), as shown in the right panel of *SI Appendix*, Fig. S13. This transition from a single hybrid state to epithelial and mesenchymal states is qualitatively the same as the results shown in Fig. 4. Thus, only one self-activation node is enough to yield the same patterns.

Validation of Bifurcated Evolution along EMT Regulatory Network. Our numerical simulation of the observed bifurcated cell-fate decision suggests that the compression-induced cell heterogeneity may be simply explained by an increase of gene-expression noise in EMT regulatory network (Fig. 4). This numerical model predicts that two genes with antagonistic interaction in EMT regulatory network are sufficient to recapitulate the bifurcated evolution of cell fate under compression by modulating the stochastic dynamics of transcription rates.

To test this prediction, we evaluate whether using only VIM and CDH1 genes can reconstruct the bifurcated cell trajectories with the mesenchymal path and epithelial path. We extract the expression levels of VIM and CDH1 from the single-cell sequencing data and plot CDH1 against VIM from the control and volumetric-compression condition. Because of the limited coverage and dropout possibility of single-cell sequencing data, we filter the data first. Practically, as we described in Fig. 3 *A–C*, there are barely any cells expressing exactly zero VIM or zero CDH1; less than 0.8% of the cells are totally negatively stained

for VIM or CDH1 probe. Based on this sc-seq data, we filter out all the cells with zero VIM or CDH1 expression as dropouts and only plot the cells with nonzero expressions (Fig. 5*A*). It is clear that cells start from a state with comparable and mild expression of VIM and CDH1 under isotonic condition (black dots in Fig. 5*A*). Cells cultured in hypertonic medium divide into two distinct regions as shown by pink dots (2.5 h) and blue dots (4 h) in Fig. 5*A*: one subgroup (mesenchymal) exhibits a high-level expression of VIM and a relatively low expression of CDH1; the other subgroup (epithelial) shows the opposite, in which the expression of VIM is low and the expression of CDH1 is high. When the expression levels of CDH1 and VIM are both low, the two regions (mesenchymal and epithelial) of cells overlap at the region where the cells originate (cells without compression). Surprisingly, the compression randomly increases the expression of one of these two genes (CDH1 and VIM), which results in the cells dividing into two distinct subgroups randomly. This observation validates that two genes with antagonistic interaction could recapitulate the bifurcated evolutions along either mesenchymal path or epithelial path. Along each path, the bimodal distribution of cells along the expression of either VIM or CDH1 is another evidence to show how the bifurcated cell-fate outcomes are generated (Fig. 5 *B* and *C*).

In addition, the sm-FISH experiment also supports the numerical simulation, in which we costain VIM and CDH1 mRNAs in single cells under different conditions of osmotic compression (Fig. 5*D*). Consistently, the sm-FISH data also recapitulates the same pattern predicated by numerical modeling, including both the mesenchymal path and the epithelial path as mentioned above (Fig. 5*E*). Moreover, the counts of transcripts of VIM and CDH1 by sm-FISH also exhibit a bimodal distribution similar to the single-cell sequencing data (Fig. 5 *F* and *G*). Overall, these quantitative results support our conclusion that the elevated gene-expression noise leads to bifurcated cell-fate outcomes following the EMT multistable network (Fig. 5*H*).

Discussion

In many types of cancer, the increased physical stress in microenvironment is a key driver of malignancy (20, 27–30). The tumor physical microenvironment is substantially remodeled from native tissue homeostasis, leading to increasing mechanical stresses (28–30), stiffened extracellular matrix (15–17), and elevated osmotic pressure (23–25). These physical stresses result in changes in cellular behaviors and functions, including modulation of integrin binding (115), activation of mechanoresponsive signaling (28, 32), cytoskeletal tension (116), TF activation (117), chromatin opening (117–119), and epigenetic modification (20, 120–123). Recent advances in single-cell technologies reveal the significant role of heterogeneity and highlight the tumor as a complex ecosystem of many species of cell types (3, 4, 124). In this paper, we report that the physical compression can induce distinct phenotypic subpopulations from its original homogeneous population of cells. The result shows that a single physical input can induce multiple differential outputs of cell fates in the same type of cancer cell, which further highlights the complexity of mechano-regulation of tumors.

Practically, we find that the newly generated subpopulations from physical compression gain the signature genes of epithelial and mesenchymal, respectively. Several recent studies have shown that the coexistence of epithelial/mesenchymal cells is important for tumor metastasis (37, 125, 126). Previous studies show that several types of physical stresses such as matrix stiffening, matrix fiber alignment, and fluid shear force can induce the EMT of the cancer cells (127, 128). Our paper shows a distinct mechanism that the physical compression promotes both the phenotypic features of epithelial and mesenchymal types. The physical compression strengthens the coexistence of epithelial cell type and mesenchymal cell type rather than deterministic generation of single phenotype as previously reported. One interesting observation is that the compression can modulate the expression of genes such as CTGF and LOX, which regulate ECM remodeling and ECM stiffening as reported (129–131). It will be interesting to study the role of CTGF in ECM remodeling and the role of LOX in ECM stiffening post osmotic compression. Thus, we can integrate several extracellular, mechanical parameters into the circuit of EMT regulatory, providing a deeper understanding of the crosstalk between the inner EMT genetic network and extracellular microenvironments. Overall, these results suggest that the physical compression may promote the evolution of tumor in an ecological way, highlighting a spectral interaction between the physical microenvironment and the ecological components of developing tumors.

Furthermore, early studies have identified EMT transition into several discrete states of its related phenotypes. With the advances in high-throughput single-cell technology, several papers employing single-cell RNA sequencing or single-cell mass cytometry have suggested that the EMT transition falls along a continuum track (11, 132–134). As reported in a more recent study, McFaline-Figueroa et al. identified continuous waves of genes regulation as opposed to discrete "partial" stages of EMT (135). To understand the discrete and continuum nature of EMT transition, this work creatively combined pooled single-cell CRISPR-Cas9 screening with single-cell sequencing. By perturbing several screened receptors, they explained how cells transiting through a continuous process appear to be in one of several discrete stages. In this paper, we integrate the concept of attractor basin into the continuum of EMT transition by employing both the single-cell transcriptome data and numerical modeling. Basically, along the continuum of the EMT transition, there are several attract basins which appear to be discrete stages in some systems; this means that there are some regions (attractor basin) in the continuum of EMT transition are stable, while other regions in the middle of the attractor basin are unstable. Practically, the cells first leave the hybrid E/M state (the first attractor basin) under compression into the continuum of the EMT regulatory topology. Eventually, the cells in the continuum of the EMT regulatory topology will fall into several other attractor basins along the continuum (E type and M type). Overall, our study supports the conclusion that there are several "checkpoints (attractor basins)" in EMT continuum that can mimic discrete stages as recently reported (135).

In addition, our work employs single-cell sequencing and sm-FISH to reveal differential changes in individual cell gene expression and found a broad agreement with prior work demonstrating large heterogeneity among a cancer population. The ability to quantitatively access the gene-expression level on a single-cell level and in a genome-wide manner highlights the power of high-throughput, unbiased single-cell technology when applied to studies of mechano-regulation, as conventional bulk assays would mask the differential effects and the stochastic nature of physical cues by averaging the expression of marker genes across the transcriptome. This work underscores the need for single-cell profiling experiments in exploring the effect of physical or mechanical cues on cancer progression and stem-cell differentiation.

Materials and Methods

Cell Lines. H1975 cell line from initial authenticated cell passages from mycoplasma were grown in Roswell Park Memorial Institute (RPMI) 1640 with 10% fetal bovine serum and 1% penicillin/streptomycin mixture.

Immunofluorescence Staining of EMT Markers. For immunofluorescence, cells were fixed in 4% paraformaldehyde, permeabilized in 0.2% Triton X-100, and

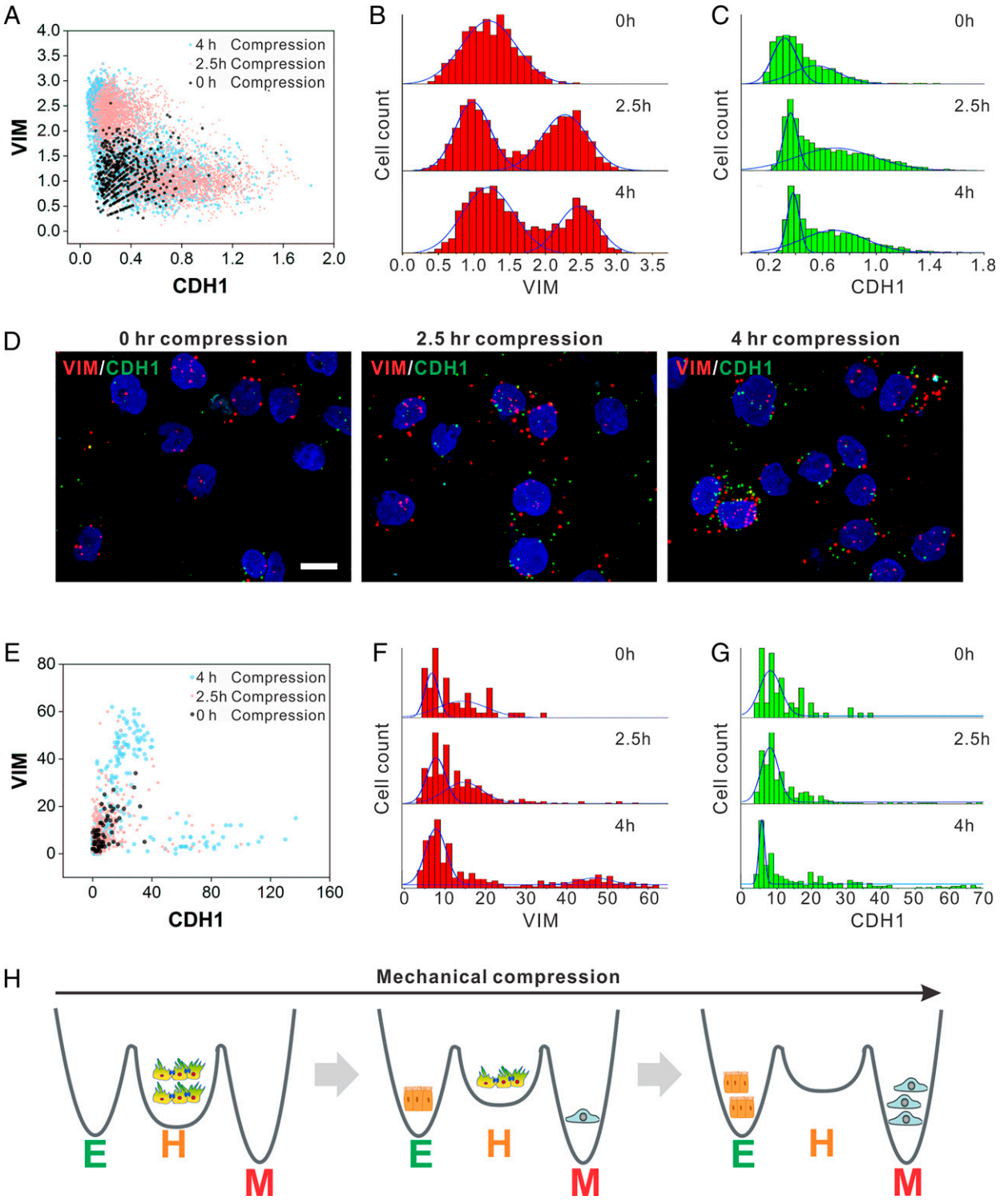

Fig. 5. Recapitulating EMT regulatory network using single-cell sequencing data and sm-FISH data, respectively. (*A*) Plot of the expression of CDH1 in each cell against its own expression of VIM, using single-cell sequencing data. (*B*) Histogram of the expression of VIM of cells at different time points before and after osmotic compression using single-cell sequencing data and indicated bimodal distribution of gene expression of VIM generated after the compression. (*C*) Histogram of the expression of CDH1 of cells at different time points before and after osmotic compression using single-cell sequencing data and indicated bimodal distribution of gene expression of CDH1 generated after the compression. (*D*) Images of sm-FISH of cells before and after osmotic compression by costaining VIM and CDH1. (Scale bar, 5 μm.) (*E*) Plot of the expression of CDH1 in each cell against its own expression of VIM, using sm-FISH data. (*F*) Histogram of the expression of VIM of cells at different time points before and after osmotic compression using sm-FISH data and indicated bimodal distribution of gene expression of VIM generated after the compression. (*G*) Histogram of the expression of CDH1 of cells at different time points before and after osmotic compression using sm-FISH data and indicated bimodal distribution of gene expression of CDH1 generated after the compression. (*H*) Metaphorical landscape illustration of our model. Each well represents a stable or metastable phenotype. The lower the well, the more stable (and the more likely to be populated) the phenotype is. Compression-induced gene-expression noise drives cells in hybrid state to deviate its original state. The cells finally generate two subpopulations in epithelial state and mesenchymal state, respectively, according to the stability of their state.

then incubated overnight with anti-E-cadherin (1:200; Catalog no. 3195, Cell Signaling Technology) and anti-Vimentin (1:200; Catalog no. ab8978, Abcam). The primary antibodies were then detected with Alexa-conjugated secondary antibodies (Life Technologies). Nuclei were visualized by costaining with DAPI.

sm-FISHs and Imaging. We performed single-molecule RNA FISH on the samples as described according to the manufacture's protocol. We costained the nuclei with DAPI to identify individual cells.

We imaged the cells with Leica SP8 confocal microscopy equipped with appropriate filter sets. The images were taken using Leica Hybrid detector. We took a series of optical z sections, each 0.35 microns high, that spanned the vertical extent of the cell.

Single-Cell RNA Sequencing. Single-cell transcriptome libraries were generated using the 10X Genomics Chromium. The libraries were firstly conversed using the MGIEasy Universal Library Conversion kit (Part No. 1000004155). After conversion step, we sequence the libraries using BGI MGISEQ-500 instrument. Single-cell transcriptome libraries were also generated using the MobiNova single-cell library construction system (MobiDrop) for validation. The libraries were sequenced with illumine sequencing platform.

Processing Single-Cell RNA Sequencing Data. A digital gene-expression matrix was constructed from the raw sequencing data using STAR algorithm in STARsolo. Data were processed using the Scanpy Python module. After quality control, H1975 cells at different time points after osmotic compression were retained. Cells in different conditions were aligned. Data were visualized using UMAP plot, similar to what was previously described. Filtered cells were clustered using Louvain clustering (Scanpy igraph method). Preprocessing steps were similar to a previous approach. In brief, genes with no count were filtered out, and each cell was normalized by the total unique molecular identifier (UMI) count per cell. The digital gene-expression matrix was renormalized after gene filtering. The data were log-transformed after adding a pseudocount and scaled to unit variance and zero mean. The dimensionality of the data was reduced by principal component analysis first and then with UMAP, followed by Louvain clustering performed on the 30 principal components (resolution = 1.5). We first fitted the top 30 principal components to compute a neighborhood graph of observations. For UMAP visualization, we directly fit the principal component analysis (PCA) matrix into the scanpy.api.tl.umap function with perplexity of 30.

For pseudotime analysis, we recalculated k-nearest neighbors at k = 15 and chose cells without osmotic compression as our initial state of cells. Pseudotime was calculated using Scanpy's PAGA function, paga. To visualize gene expression in pseudotime, cells were ordered according to their pseudotime values from least to greatest. The expression level of three epithelial genes and three mesenchymal genes were plotted according to the pseudotime ordering.

Osmotic Stress. Hypertonic stress was applied by adding PEG 300 to isotonic culture medium. The correlation between molarity and osmolality for solutions of PEG 300 in water was obtained from previous measurement (136). The actual osmotic pressure applied to cells was calculated by summing up the calculated osmolality of PEG and isotonic medium (325 mOsm) and was further validated by measurement using a micro-osmometer (model 3300, Advanced Instruments, Inc.). Cells were incubated for 10 min at 37 °C and 5% CO_2 for equilibrate in PEG solution. The cell size and mechanics achieved equilibration within 2 min after adding PEG 300 based on previous studies (5,6).

Quantification and Statistical Details. Statistically significant differences between the means of two groups were assessed using a Student's *t* test, whereas data containing more than two experimental groups were analyzed with a one-way ANOVA followed by a Bonferroni's multiple-comparison test. In the experiment of sucrose gradient assay, the statistically significant differences are evaluated using multivariate ANOVA. Error bars in all figures represent SD. All statistical analyses were performed in the Origin 9.0 software. $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$, respectively.

Data Availability. All the raw sequencing data have been deposited at National Center for Biotechnology Information Sequence Read Archive (BioProject ID: PRJNA770977). All other study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. M.G. acknowledges the support from National Institute of General Medical Sciences Grant No. 1R01GM140108, the Jeptha H. and Emily V. Wade Award at Massachusetts Institute of Technology, and the Sloan Research Fellowship. Y.L. acknowledges the financial support from the National Natural Science Foundation of China (Grant Nos. 32171248 and 12102142) and the Fundamental Research Funds for the Central Universities, Huazhong University of Science and Technology (2021GCRC056). Y.L. thanks Dr. Wenshan Zheng (MobiDrop, Inc.) for helpful discussion and technical support.

1. R. A. Burrell, N. McGranahan, J. Bartek, C. Swanton, The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338–345 (2013).
2. P. B. Gupta *et al.*, Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011).
3. J. Peng *et al.*, Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res.* **29**, 725–738 (2019).
4. S. F. Roerink *et al.*, Intra-tumour diversification in colorectal cancer at the single-cell level. *Nature* **556**, 457–462 (2018).
5. J. Irianto *et al.*, DNA damage follows repair factor depletion and portends genome variation in cancer cells after pore migration. *Curr. Biol.* **27**, 210–223 (2017).
6. T. Takaki *et al.*, Actomyosin drives cancer cell nuclear dysmorphia and threatens genome stability. *Nat. Commun.* **8**, 16013 (2017).
7. R. P. das Neves *et al.*, Connecting variability in global transcription rate to mitochondrial variability. *PLoS Biol.* **8**, e1000560 (2010).
8. A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, S. Tyagi, Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309 (2006).
9. A. K. Shalek *et al.*, Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* **498**, 236–240 (2013).
10. J. D. Buenrostro *et al.*, Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486–490 (2015).
11. A. P. Deshmukh *et al.*, Identification of EMT signaling cross-talk and gene regulatory networks by single-cell RNA sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102050118 (2021).
12. A. K. Shalek *et al.*, Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363–369 (2014).
13. M. Darnell, L. Gu, D. Mooney, RNA-seq reveals diverse effects of substrate stiffness on mesenchymal stem cells. *Biomaterials* **181**, 182–188 (2018).
14. M. Darnell *et al.*, Material microenvironmental properties couple to induce distinct transcriptional programs in mammalian stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8368–E8377 (2018).
15. K. R. Levental *et al.*, Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* **139**, 891–906 (2009).
16. B. R. Seo *et al.*, Obesity-dependent changes in interstitial ECM mechanics promote breast tumorigenesis. *Sci. Transl. Med.* **7**, 301ra130 (2015).
17. S. C. Wei *et al.*, Matrix stiffness drives epithelial-mesenchymal transition and tumour metastasis through a TWIST1-G3BP2 mechanotransduction pathway. *Nat. Cell Biol.* **17**, 678–688 (2015).
18. D. Jia *et al.*, Distinguishing mechanisms underlying EMT tristability. *Cancer Converg.* **1**, 2 (2017).
19. G. Higgins *et al.*, Decreased cell stiffness facilitates cell detachment and cell migration from breast cancer spheroids in 3D collagen matrices of different rigidity. bioRxiv [Preprint] (2021). https://doi.org/10.1101/2021.01.21.427639 (Accessed 21 January 2021).
20. R. S. Stowers *et al.*, Matrix stiffness induces a tumorigenic phenotype in mammary epithelium through changes in chromatin accessibility. *Nat. Biomed. Eng.* **3**, 1009–1019 (2019).
21. Y. Li, W. Tang, M. Guo, The cell as matter: Connecting molecular biology to cellular functions. *Matter* **4**, 1863–1891 (2021).
22. X. Zhao *et al.*, Rapid generation of hybrid biochemical/mechanical cues in heterogeneous droplets for high-throughput screening of cellular response. *Lab Chip* **21**, 2691–2701 (2021).
23. R. Gutmann *et al.*, Interstitial hypertension in head and neck tumors in patients: Correlation with tumor size. *Cancer Res.* **52**, 1993–1995 (1992).
24. C.-H. Heldin, K. Rubin, K. Pietras, A. Östman, High interstitial fluid pressure - An obstacle in cancer therapy. *Nat. Rev. Cancer* **4**, 806–813 (2004).
25. P. A. Netti, L. T. Baxter, Y. Boucher, R. Skalak, R. K. Jain, Time-dependent behavior of interstitial fluid pressure in solid tumors: Implications for drug delivery. *Cancer Res.* **55**, 5451–5458 (1995).
26. Y. Li *et al.*, Compression-induced dedifferentiation of adipocytes promotes tumor progression. *Sci. Adv.* **6**, eaax5611 (2020).
27. Y. L. Han *et al.*, Cell swelling, softening and invasion in a three-dimensional breast cancer model. *Nat. Phys.* **16**, 101–108 (2020).
28. M. E. Fernández-Sánchez *et al.*, Mechanical induction of the tumorigenic β-catenin pathway by tumour growth pressure. *Nature* **523**, 92–95 (2015).
29. J. M. Tse *et al.*, Mechanical compression drives cancer cells toward invasive phenotype. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 911–916 (2012).
30. H. T. Nia *et al.*, Solid stress and elastic energy as measures of tumour mechanopathology. *Nat. Biomed. Eng.* **1**, 0004 (2016).
31. F. Spill, D. S. Reynolds, R. D. Kamm, M. H. Zaman, Impact of the physical microenvironment on tumor progression and metastasis. *Curr. Opin. Biotechnol.* **40**, 41–48 (2016).
32. T. Panciera, L. Azzolin, M. Cordenonsi, S. Piccolo, Mechanobiology of YAP and TAZ in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **18**, 758–770 (2017).
33. Y. Li *et al.*, Helical nanofiber yarn enabling highly stretchable engineered microtissue. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9245–9250 (2019).

34. F. Spill, C. Bakal, M. Mak, Mechanical and systems biology of cancer. *Comput. Struct. Biotechnol. J.* **16**, 237–245 (2018).
35. X. Zhang, M. Mak, Biophysical informatics approach for quantifying phenotypic heterogeneity in cancer cell migration in confined microenvironments. *Bioinformatics* **37**, btab053 (2021).
36. S. A. Sevier, D. A. Kessler, H. Levine, Mechanical bounds to transcriptional noise. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13983–13988 (2016).
37. I. Pastushenko *et al.*, Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463–468 (2018).
38. I. Y. Wong *et al.*, Collective and individual migration following the epithelial-mesenchymal transition. *Nat. Mater.* **13**, 1063–1071 (2014).
39. J. T. George, M. K. Jolly, S. Xu, J. A. Somarelli, H. Levine, Survival outcomes in cancer patients predicted by a partial EMT gene expression scoring metric. *Cancer Res.* **77**, 6415–6428 (2017).
40. D. Jia *et al.*, Testing the gene expression classification of the EMT spectrum. *Phys. Biol.* **16**, 025002 (2019).
41. D. L. Roden *et al.*, Single cell transcriptomics reveals molecular subtype and functional heterogeneity in models of breast cancer. bioRxiv [Preprint] (2018). https://doi.org/10.1101/282079 (Accessed 14 March 2018).
42. L. Cao *et al.*, OCT4 increases BIRC5 and CCND1 expression and promotes cancer progression in hepatocellular carcinoma. *BMC Cancer* **13**, 82 (2013).
43. H. Okuda *et al.*, miR-7 suppresses brain metastasis of breast cancer stem-like cells by modulating KLF4. *Cancer Res.* **73**, 1434–1444 (2013).
44. Y.-Y. Hou *et al.*, MicroRNA-519d targets MKi67 and suppresses cell growth in the hepatocellular carcinoma cell line QGY-7703. *Cancer Lett.* **307**, 182–190 (2011).
45. S. Tripathi, P. Chakraborty, H. Levine, M. K. Jolly, A mechanism for epithelial-mesenchymal heterogeneity in a population of cancer cells. *PLOS Comput. Biol.* **16**, e1007619 (2020).
46. J. Yang *et al.*, EMT International Association (TEMTIA), Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **21**, 341–352 (2020).
47. F. A. Wolf *et al.*, PAGA: Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).
48. T. M. Gierahn *et al.*, Seq-well: Portable, low-cost RNA sequencing of single cells at high throughput. *Nat. Methods* **14**, 395–398 (2017).
49. G. X. Zheng *et al.*, Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
50. E. Z. Macosko *et al.*, Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
51. A. M. Klein *et al.*, Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
52. Z. Chen, C. M. Fillmore, P. S. Hammerman, C. F. Kim, K.-K. Wong, Non-small-cell lung cancers: A heterogeneous set of diseases. *Nat. Rev. Cancer* **14**, 535–546 (2014).
53. M. A. Friedman, E. Slater, Malignant pleural effusions. *Cancer Treat. Rev.* **5**, 49–66 (1978).
54. A. G. Henderson *et al.*, Cystic fibrosis airway secretions exhibit mucin hyperconcentration and increased osmotic pressure. *J. Clin. Invest.* **124**, 3047–3060 (2014).
55. P. L. Luz da, H. Shubin, M. H. Weil, E. Jacobson, L. Stein, Pulmonary edema related to changes in colloid osmotic and pulmonary artery wedge pressure in patients after acute myocardial infarction. *Circulation* **51**, 350–357 (1975).
56. L. R. Schiller, M. Emmett, C. A. Santa Ana, J. S. Fordtran, Osmotic effects of polyethylene glycol. *Gastroenterology* **94**, 933–941 (1988).
57. B. Akabayov, S. R. Akabayov, S.-J. Lee, G. Wagner, C. C. Richardson, Impact of macromolecular crowding on DNA replication. *Nat. Commun.* **4**, 1615 (2013).
58. E. Becht *et al.*, Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38 (2018).
59. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
60. M. K. Jolly *et al.*, Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget* **7**, 27067–27084 (2016).
61. L. Kester, A. van Oudenaarden, Single-cell transcriptomics meets lineage tracing. *Cell Stem Cell* **23**, 166–179 (2018).
62. M. Plass *et al.*, Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science* **360**, eaaq1723 (2018).
63. W. Saelens, R. Cannoodt, H. Todorov, Y. Saeys, A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* **37**, 547–554 (2019).
64. F. Bocci *et al.*, NRF2 activates a partial epithelial-mesenchymal transition and is maximally present in a hybrid epithelial/mesenchymal phenotype. *Integr. Biol.* **11**, 251–263 (2019).
65. P. Chakraborty, J. T. George, S. Tripathi, H. Levine, M. K. Jolly, Comparative study of transcriptomics-based scoring metrics for the epithelial-hybrid-mesenchymal spectrum. *Front. Bioeng. Biotechnol.* **8**, 220 (2020).
66. X.-J. Tian, H. Zhang, J. Xing, Coupled reversible and irreversible bistable switches underlying TGFβ-induced epithelial to mesenchymal transition. *Biophys. J.* **105**, 1079–1089 (2013).
67. J. Zhang *et al.*, TGF-β-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci. Signal.* **7**, ra91–ra91 (2014).
68. M. Zeisberg, A. A. Shah, R. Kalluri, Bone morphogenic protein-7 induces mesenchymal to epithelial transition in adult renal fibroblasts and facilitates regeneration of injured kidney. *J. Biol. Chem.* **280**, 8094–8100 (2005).
69. J. Hasty, J. Pradines, M. Dolnik, J. J. Collins, Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2075–2080 (2000).
70. H. Maamar, A. Raj, D. Dubnau, Noise in gene expression determines cell fate in Bacillus subtilis. *Science* **317**, 526–529 (2007).
71. J. M. Raser, E. K. O'Shea, Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811–1814 (2004).
72. A. Sanchez, I. Golding, Genetic determinants and cellular constraints in noisy gene expression. *Science* **342**, 1188–1193 (2013).
73. J. Stewart-Ornstein, J. S. Weissman, H. El-Samad, Cellular noise regulons underlie fluctuations in Saccharomyces cerevisiae. *Mol. Cell* **45**, 483–493 (2012).
74. A. Baudrimont, V. Jaquet, S. Wallerich, S. Voegeli, A. Becskei, Contribution of RNA degradation to intrinsic and extrinsic noise in gene expression. *Cell Rep.* **26**, 3752–3761.e3755 (2019).
75. H. H. Chang, M. Hemberg, M. Barahona, D. E. Ingber, S. Huang, Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**, 544–547 (2008).
76. H. Kempe, A. Schwabe, F. Crémazy, P. J. Verschure, F. J. Bruggeman, The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise. *Mol. Biol. Cell* **26**, 797–804 (2015).
77. S. C. Little, M. Tikhonov, T. Gregor, Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell* **154**, 789–800 (2013).
78. K. Miller-Jensen, S. S. Dey, D. V. Schaffer, A. P. Arkin, Varying virulence: Epigenetic control of expression noise and disease processes. *Trends Biotechnol.* **29**, 517–525 (2011).
79. O. Padovan-Merhar *et al.*, Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Mol. Cell* **58**, 339–352 (2015).
80. G. Ren *et al.*, CTCF-mediated enhancer-promoter interaction is a critical regulator of cell-to-cell variation of gene expression. *Mol. Cell* **67**, 1049–1058.e1046 (2017).
81. L. Wang *et al.*, Bistable switches control memory and plasticity in cellular differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6638–6643 (2009).
82. T. S. Gardner, C. R. Cantor, J. J. Collins, Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339–342 (2000).
83. N. D. Lord *et al.*, Stochastic antagonism between two proteins governs a bacterial cell fate switch. *Science* **366**, 116–120 (2019).
84. B. Munsky, G. Neuert, A. van Oudenaarden, Using gene expression noise to understand gene regulation. *Science* **336**, 183–187 (2012).
85. S. Sarkar, S. K. Sinha, H. Levine, M. K. Jolly, P. S. Dutta, Anticipating critical transitions in epithelial-hybrid-mesenchymal cell-fate determination. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 26343–26352 (2019).
86. M. Acar, J. T. Mettetal, A. van Oudenaarden, Stochastic switching as a survival strategy in fluctuating environments. *Nat. Genet.* **40**, 471–475 (2008).
87. E. Kussell, S. Leibler, Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**, 2075–2078 (2005).
88. I. Hecht, D. A. Kessler, H. Levine, Transient localized patterns in noise-driven reaction-diffusion systems. *Phys. Rev. Lett.* **104**, 158301 (2010).
89. B. Hu, D. A. Kessler, W.-J. Rappel, H. Levine, Effects of input noise on a simple biochemical switch. *Phys. Rev. Lett.* **107**, 148101 (2011).
90. K. Wang, W.-J. Rappel, R. Kerr, H. Levine, Quantifying noise levels of intercellular signals. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **75**, 061905 (2007).
91. M. Guo *et al.*, Cell volume change through water efflux impacts cell stiffness and stem cell fate. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8618–E8627 (2017).
92. Y. Li *et al.*, Volumetric compression induces intracellular crowding to control intestinal organoid growth via Wnt/β-catenin signaling. *Cell Stem Cell* **28**, 63–78.e67 (2021).
93. E. H. Zhou *et al.*, Universal behavior of the osmotically compressed cell and its analogy to the colloidal glass transition. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10632–10637 (2009).
94. N. Battich, T. Stoeger, L. Pelkmans, Control of transcript variability in single mammalian cells. *Cell* **163**, 1596–1610 (2015).
95. D. Grün, L. Kester, A. van Oudenaarden, Validation of noise models for single-cell transcriptomics. *Nat. Methods* **11**, 637–640 (2014).
96. M. Mak, F. Spill, R. D. Kamm, M. H. Zaman, Single-cell migration in complex microenvironments: Mechanics and signaling dynamics. *J. Biomech. Eng.* **138**, 021004 (2016).
97. S. Tripathi, H. Levine, M. K. Jolly, The physics of cellular decision making during epithelial–mesenchymal transition. *Annu. Rev. Biophys.* **49**, 1–18 (2020).
98. X. Zhang, T. Chan, M. Mak, Morphodynamic signatures of MDA-MB-231 single cells and cell doublets undergoing invasion in confined microenvironments. *Sci. Rep.* **11**, 6529 (2021).
99. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
100. M. M. Hansen *et al.*, Macromolecular crowding creates heterogeneous environments of gene expression in picolitre droplets. *Nat. Nanotechnol.* **11**, 191–197 (2016).
101. B. De Craene, G. Berx, Regulatory networks defining EMT during cancer initiation and progression. *Nat. Rev. Cancer* **13**, 97–110 (2013).
102. F. Font-Clos, S. Zapperi, C. A. M. La Porta, Topography of epithelial-mesenchymal plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 5902–5907 (2018).
103. M. Lu, M. K. Jolly, H. Levine, J. N. Onuchic, E. Ben-Jacob, MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18144–18149 (2013).

104. S. E. Leggett, A. M. Hruska, M. Guo, I. Y. Wong, The epithelial-mesenchymal transition and the cytoskeleton in bioengineered systems. *Cell Commun. Signal.* **19**, 32 (2021).
105. S. Tripathi, J. Xing, H. Levine, M. K. Jolly, "Mathematical modeling of plasticity and heterogeneity in EMT" in *The Epithelial-to Mesenchymal Transition*, K. Campbell, E. Theveneau, Eds. (Methods in Molecular Biology, Springer, 2021), vol. 2179, pp. 385–413.
106. T. G. Graham, S. M. Tabei, A. R. Dinner, I. Rebay, Modeling bistable cell-fate choices in the Drosophila eye: Qualitative and quantitative perspectives. *Development* **137**, 2265–2278 (2010).
107. N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, M. B. Elowitz, Gene regulation at the single-cell level. *Science* **307**, 1962–1965 (2005).
108. Y. Setty, A. E. Mayo, M. G. Surette, U. Alon, Detailed map of a cis-regulatory input function. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7702–7707 (2003).
109. A. Duncan, S. Liao, T. Vejchodský, R. Erban, R. Grima, Noise-induced multistability in chemical systems: Discrete versus continuum modeling. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **91**, 042111 (2015).
110. T. Gao, J. Duan, X. Li, Fokker–Planck equations for stochastic dynamical systems with symmetric Lévy motions. *Appl. Math. Comput.* **278**, 1–20 (2016).
111. H. Risken, "Fokker-Planck equation" in *The Fokker-Planck Equation* (Springer, 1996), pp. 63–95.
112. N. Kumar, T. Platini, R. V. Kulkarni, Exact distributions for stochastic gene expression models with bursting and feedback. *Phys. Rev. Lett.* **113**, 268105 (2014).
113. A. Raj, A. van Oudenaarden, Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008).
114. V. Shahrezaei, P. S. Swain, Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17256–17261 (2008).
115. N. B. M. Rafiq *et al.*, A mechano-signalling network linking microtubules, myosin IIA filaments and integrin-based adhesions. *Nat. Mater.* **18**, 638–649 (2019).
116. D. A. Fletcher, R. D. Mullins, Cell mechanics and the cytoskeleton. *Nature* **463**, 485–492 (2010).
117. D. Cai *et al.*, Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression. *Nat. Cell Biol.* **21**, 1578–1589 (2019).
118. K. Damodaran *et al.*, Compressive force induces reversible chromatin condensation and cell geometry-dependent transcriptional response. *Mol. Biol. Cell* **29**, 3039–3051 (2018).
119. A. Tajik *et al.*, Transcription upregulation via force-induced direct stretching of chromatin. *Nat. Mater.* **15**, 1287–1296 (2016).
120. S. W. Crowder, V. Leonardo, T. Whittaker, P. Papathanasiou, M. M. Stevens, Material cues as potent regulators of epigenetics and stem cell function. *Cell Stem Cell* **18**, 39–52 (2016).
121. F. Alisafaei, D. S. Jokhun, G. V. Shivashankar, V. B. Shenoy, Regulation of nuclear architecture, mechanics, and nucleocytoplasmic shuttling of epigenetic factors by cell geometric constraints. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 13200–13209 (2019).
122. W. Jia, A. Deshmukh, S. A. Mani, M. K. Jolly, H. Levine, A possible role for epigenetic feedback regulation in the dynamics of the epithelial-mesenchymal transition (EMT). *Phys. Biol.* **16**, 066004 (2019).
123. W. Jia *et al.*, Epigenetic feedback and stochastic partitioning during cell division can drive resistance to EMT. *Oncotarget* **11**, 2611–2624 (2020).
124. A. R. Lourenco *et al.*, Differential contributions of pre-and post-EMT tumor cells in breast cancer metastasis. *Cancer Res.* **80**, 163–169 (2020).
125. V. Bernard *et al.*, Single-cell transcriptomics of pancreatic cancer precursors demonstrates epithelial and microenvironmental heterogeneity as an early event in neoplastic progression. *Clin. Cancer Res.* **25**, 2194–2205 (2019).
126. S. A. Mani *et al.*, The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715 (2008).
127. R. L. Heise, V. Stober, C. Cheluvaraju, J. W. Hollingsworth, S. Garantziotis, Mechanical stretch induces epithelial-mesenchymal transition in alveolar epithelia via hyaluronan activation of innate immunity. *J. Biol. Chem.* **286**, 17435–17444 (2011).
128. L. Przybyla, J. M. Muncie, V. M. Weaver, Mechanical control of epithelial-to-mesenchymal transitions in development and cancer. *Annu. Rev. Cell Dev. Biol.* **32**, 527–554 (2016).
129. Y. Deng, P. Chakraborty, M. K. Jolly, H. Levine, A theoretical approach to coupling the epithelial-mesenchymal transition (EMT) to extracellular matrix (ECM) stiffness via LOXL2. *Cancers (Basel)* **13**, 1609 (2021).
130. B. Wei *et al.*, Human colorectal cancer progression correlates with LOX-induced ECM stiffening. *Int. J. Biol. Sci.* **13**, 1450–1457 (2017).
131. J. M. Petrosino, A. Leask, F. Accornero, Genetic manipulation of CCN2/CTGF unveils cell-specific ECM-remodeling effects in injured skeletal muscle. *FASEB J.* **33**, 2047–2057 (2019).
132. P. Chakraborty, J. T. George, W. A. Woodward, H. Levine, M. K. Jolly, Gene expression profiles of inflammatory breast cancer reveal high heterogeneity across the epithelial-hybrid-mesenchymal spectrum. *Transl. Oncol.* **14**, 101026 (2021).
133. M. Mandal *et al.*, Modeling continuum of epithelial mesenchymal transition plasticity. *Integr. Biol.* **8**, 167–176 (2016).
134. K. P. Simeonov *et al.*, Single-cell lineage tracing of metastatic cancer reveals selection of hybrid EMT states. *Cancer Cell* **39**, 1150–1162.e9 (2021).
135. J. L. McFaline-Figueroa *et al.*, A pooled single-cell genetic screen identifies regulatory checkpoints in the continuum of the epithelial-to-mesenchymal transition. *Nat. Genet.* **51**, 1389–1398 (2019).
136. N. P. Money, Osmotic pressure of aqueous polyethylene glycols: Relationship between molecular weight and vapor pressure deficit. *Plant Physiol.* **91**, 766–769 (1989).