

Report

Deformability Cytometry: Unraveling the Impact of Gene ME480 on Cell Mechanics through High-Speed Imaging and Automated Analysis

Yung-Cheng Chiang¹, Clarisse Jeanne Maylis Pierre¹, Antoine Edouard Chekireb-Pack¹, Nadim Daher², and Michel Chammas²

¹School of Life Sciences, École polytechnique fédérale de Lausanne

²College of Management of Technology, École polytechnique fédérale de Lausanne

ABSTRACT

Mechanical phenotyping of single cells has long been a less-explored territory in the field of mechanobiology, given the difficulty of probing cells without perturbing their viability. Label-free and fluorescence methods have been proposed and validated in settings such as cell differentiation, cancer metastasis, and immunology. However, these techniques generally have throughput rates of 10-100 cells per hour, which is way too low for profiling cell populations with statistical significance. Here, we demonstrated how deformability cytometry (DC) can be a promising solution. By implementing the analytical part of DC, we successfully extract mechanical features for phenotyping cell lines in response to ME480 protein expression. In addition, we discussed potential confounders that could bias the experiment, highlighting the strengths and drawbacks of the methods.

INTRODUCTION

The field of mechanobiology, which explores the interplay between mechanical forces and biological processes, has emerged as a cornerstone in understanding the complexities of life. It offers crucial insights into how physical forces and material properties govern cellular behaviors, impacting a wide range of biological contexts from development to disease pathogenesis. Notably, the study of mechanobiology has illuminated the mechanisms underlying various pathologies, including cancer, where mechanical deregulation often plays a pivotal role (1). Central to this exploration is the concept of cell deformation - a key factor in vital processes such as migration, mitosis, and the formation of complex three-dimensional structures. Understanding how cells deform and respond to mechanical stimuli is essential for a comprehensive grasp of their function and behavior in physiological and pathological states (2).

In this context, real-time deformability cytometry (RT-DC) emerges as a powerful tool, allowing for the high-throughput assessment of cellular mechanical properties. Proposed by Gossett et al. and improved by Otto et al., RT-DC revolutionizes the ability to measure cell deformability in real-time under physiologically relevant conditions, providing a window into the cellular mechanical phenotype (1, 3). As **Figure 1** shows, cells are injected and flowed through a PDMS microfluidic channel constriction where cells are deformed without contact by shear stresses and pressure gradients. Deformed cells are illuminated with an LED and imaged in real-time with a CMOS camera at a very high frame rate (1). The deformation

of the cell would be evaluated together with other parameters, either through multidimensional visualization, statistical tests, or hydrodynamic model simulation. With its ability to rapidly sample large cell populations with quantifiable mechanical metrics, the label-free method is invaluable in research areas such as cancer biology and immunobiology, where rapid and high-throughput readout of cell type signature can bring great biomedical benefits.

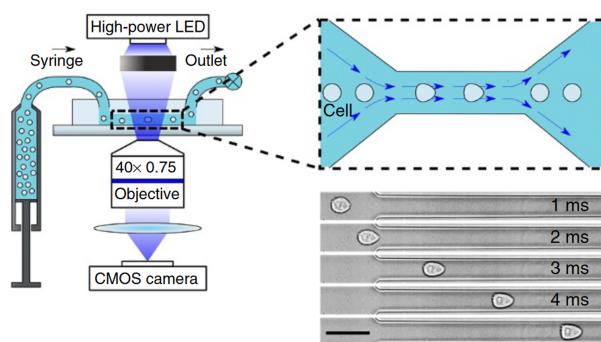


Figure 1: Real-time deformability cytometry setup and measurement principle. Scale bar, 50μm (Figure revised from Otto et al. 2015 (1))

In this study, we leveraged the capabilities of DC to investigate the mechanical implications of ME480 expression on cellular deformability. A comprehensive, image-processing-based analytical pipeline for cell detection, segmentation, and

quantification was developed. Firstly, a statistical study was done to identify sample detachment time as a confounding factor for DC measurement. Secondly, a comparative analysis between cells with wild-type and modified ME480 expression was performed. We discovered that ME480 yields higher cell deformation, which implies its potential role in the regulation of cellular structures. Moreover, we identify the cell-dependent ME480 effect, which reflects on the change in the size of cells. Selected intensity metrics were included as deterministic features for separating cell populations. We also discussed other confounders, hypotheses, and workflow details that could be part of the future work.

MATERIALS AND METHODS

Microfluidic experiments

Cells in suspension are centrifuged and resuspended in a solution of phosphate saline buffer (without Mg^{2+} and Ca^{2+}) containing methylcellulose to a final concentration of 10^6 cells/ml at $37^\circ C$. The addition of methylcellulose increases the density of the buffer and thus reduces sedimentation of the cells during the experiment. In addition, it also increases the viscosity of the medium to allow higher shear forces at lower flow velocity (1, 2).

The cell suspension is then drawn into a syringe that is connected to the assembled microfluidic chip by polymer tubing which is extensively cleaned. Before starting any measurement, the flow is stabilized at $4\mu L/min$. Every data acquisition is usually carried out at the rear part of the $300\mu m$ -long constriction where the deformation and thus the cell shape will have reached a steady state. The microfluidic plate is placed under the Phantom VE 640 L camera that performs image acquisition with parameters listed in **Table 2**. Cells are then sent down the channel with a sheath flow rate of $12\mu L/min$ to stabilize them and allow them to be in focus for our acquisition (1, 2). Notably, instead of streamlining the analytical pipeline with the image acquisition step, here we did the image analysis afterward (i.e., DC instead of RT-DC).

Table 1: Flow Cytometry Parameters

Cell sample flow rate	$4\mu L/min$
Sheath flow rate	$12\mu L/min$
Channel dimensions (length,width,depth)	$300 \times 30 \times 30 \mu m$

Image processing

Image stack size reduction

A script supported by Python OpenCV (4), scikit-image (5) was applied to identify and remove frames without the presence of cells. The raw stack was downsampled 64 times (8 times on x and y dimensions) to save memory. A background derived from the median of 100 random frames was subtracted from the stack. Then variance of each frame was computed. As

Table 2: Image Acquisition Parameters

Camera	Phantom VE 640 L
Image type	Phase contrast
Exposure time	$1\mu s$
Frame rate	10,000 fps
Number of frames per movie	356,953
Pixel size	$10 \times 10 \mu m$
Magnification	x30
Image dimensions	256×128 pixels

shown in **Figure 2**, the presence of the cells in the microfluidics was highlighted with a clear variance peak. With such a method, 27252 among 356953 frames were characterized as "cell presenting" (13.1x reduction ratio). All image processing steps could be easily done in 5 minutes (See Part A in **Table 6**).

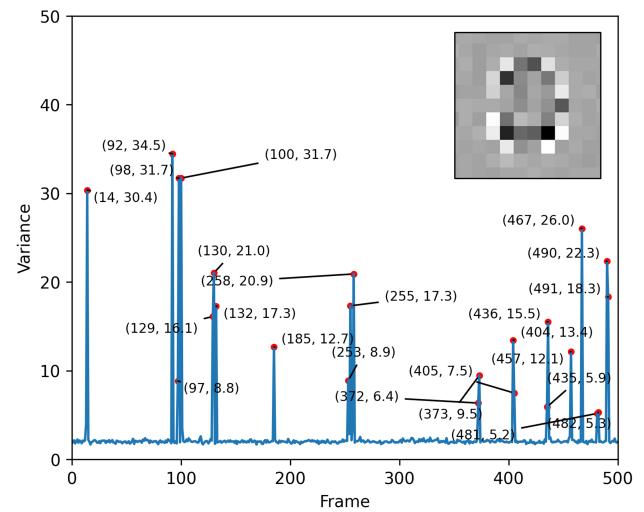


Figure 2: Size reduction based on downsampling and variance thresholding. Variance of first 500 background subtracted frames. Here peaks with values larger than 5 were labelled. (inlet) The downsampled cell image.

Single-cell segmentation and quantification

A concatenated image processing pipeline powered by Python OpenCV (4), scikit-image (5), and ImageJ scripts was used to quantify individual cells in high throughput. For cellular segmentation, the main idea is to capture the thin bright halo around the cells caused by slight under-focus (6). Firstly, a median background was subtracted from the stack. Secondly, a difference of Gaussian (DoG, $\sigma_1 = 1$, $\sigma_2 = 5$) bandpass filter followed by intensity thresholding was applied to extract the cell contour. Then, a series of morphological operations (i.e., small object removal and erosion) were done sequentially to obtain the binary mask representing the single cell.

With commercially available Intel Core i7 Processors, the

Python code managed to segment >120 frames per second. As shown in **Figure 3**, the algorithm successfully segmented around 90% of the samples. Setting a wider DoG filter could yield a higher percentage of capture but with the expense of more irregular contours impacted by the over-blurred bright halo. Here we keep a narrow DoG for better quantification of individual cells. In addition, the algorithm also wrongly segmented enormous cellular compartments clogging the channel (not shown here). These huge parts were considered outliers and would be thus ruled out from the downstream analysis.

For measurement, the binary objects with sizes larger than $36\mu\text{m}^2$ (i.e., 400 pixels) were captured. Objects touching the edge of the image frame were discarded due to the potential quantification bias. To extract the grayscale pixel intensity profile of the image stack, the measurement of regions of interest (ROIs) was done by overlaying binary masks onto the raw image stack (See Part C in **Table 6**).

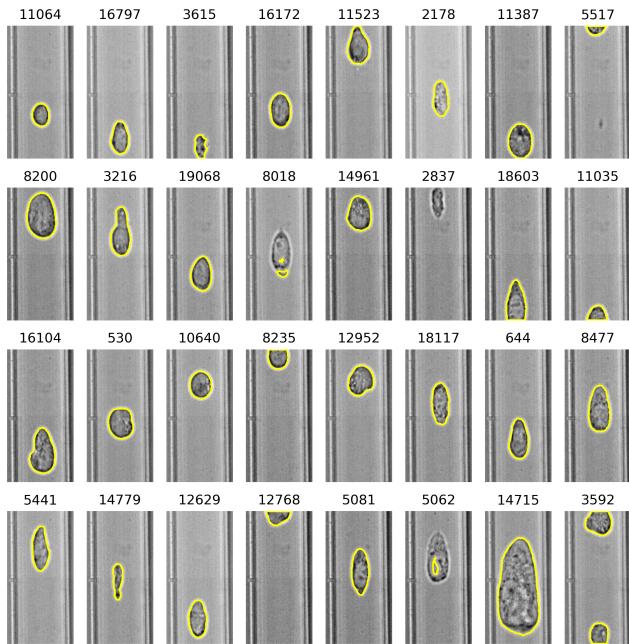


Figure 3: Randomly selected frames from CellA_GFP overlaid with segmented contours (indicated with yellow lines).

Data analysis and statistical methods

A Python-centered workflow was applied. Data points were filtered based on cross-sectional area, aspect ratio, and area ratio (i.e., the inverse of solidity) to discard cell debris, doublets, and concave objects (2, 6).

By the nature of the experimental setup, three distinct types of mechanical readouts are available: size, morphology, and intensity profile. In addition to available metrics accessed in ImageJ, deformation is defined as $D = 1 - \text{Solidity}$. A quality control step was done by plotting selected metrics

Table 3: **Filter Parameters**

Parameter	Min	Max
Area (μm^2)	36	400
Aspect ratio	0.5	2
Area ratio (1/Solidity)	1	1.05

over slices and performing linear fit. As shown in **Figure 4**, none of the metrics show serious shifts throughout the flow experiment. It implies that the effect of flowing order can be neglected. The time consistency of the measurement stays true. If there is any statistical difference between different experimental groups, it might be due to other factors (i.e., treatments, batch effect...).

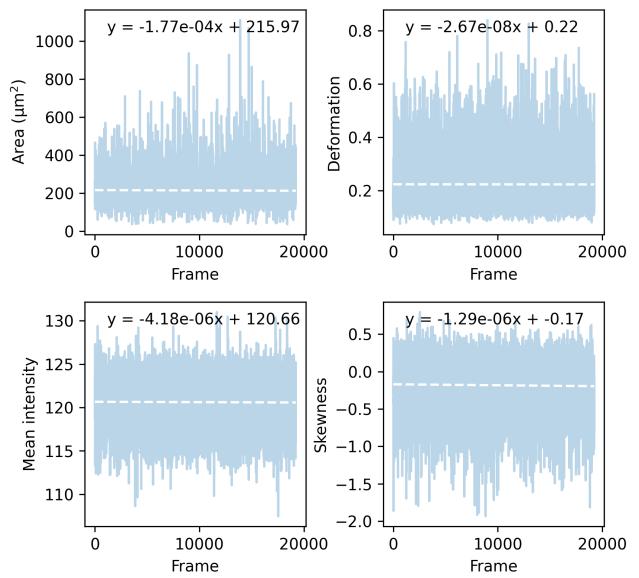


Figure 4: Quality control of CellA_GFP. The horizontal axis indicates the time point when the given sample was flown into the microfluidic chip. Linear equation was fit and plotted as white lines.

Statistical tests using Python scikit-learn (7), statannotations (8), and statsmodels (9) (i.e., Welch t-test, linear mixed effect models) were applied to study if cell lines or genetic treatments have a significant effect on cellular properties. Visualization was done either with a multi-dimensional scatter plot or T-distributed neighbor embedding (tSNE) supported by Matplotlib (10) and seaborn (11) (See Part B and C in **Table 6**).

RESULTS

Detachment time introduces incertitude in deformation cytometry measurements

DC sample preparation and measurement take time, and the differential delay might play a confounding role. When

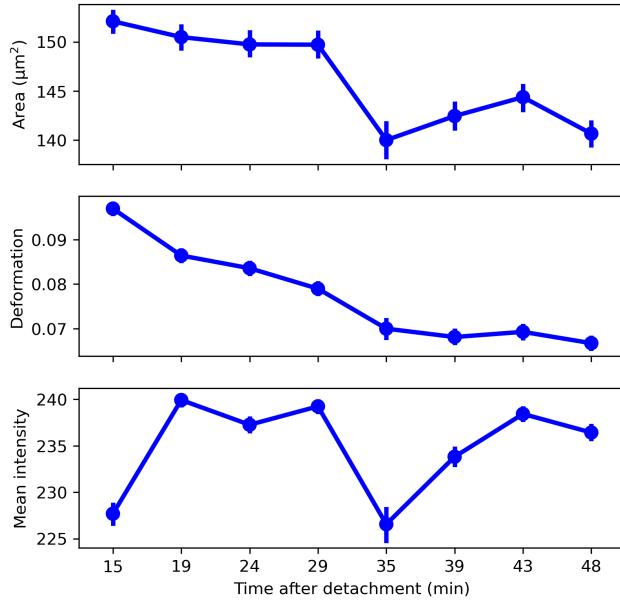


Figure 5: CellA_ShME480 metrics measured at different time points after detachment.

comparing the CellA_ShME480 samples recorded at different time points after cell detachment from the petri dish, it is found that among all types of measurements, size and morphological metrics were significantly decreasing as the time delay went larger, while the intensity profile fluctuated, as shown in **Figure 5**. This indicates that CellA_ShME480 became smaller and rounder after a longer time delay. One possible explanation is that the transition from cell-cell or cell-surface contact to a single-cell state may result in mechanical adaptation associated with F-actin remodeling (12). Such rearrangement of the cytoskeleton is a dynamic process and can be cell-type specific. To avoid propagating the uncertainty of detachment time to subsequent measurement and hypothesis testing, it is crucial to optimize the cell preparation step by aligning all detachment time for every experimental condition. Otherwise, the comparison of obtained metrics becomes not reliable.

ME480 facilitates cell-type dependent deformation dynamics

We then study the impact of ME480 protein on cellular mechanics. As shown in **Table 4**, 4 cell lines derived from 2 cell types were investigated. The expression of transfected genes is designed to be activated by exposure to doxycycline. CellA_GFP and CellB_ME480 express either endogenous or exogenous ME480 proteins, whereas CellA_shM480 and CellB_GFP have no expression. If ME480 influences the cellular mechanics, we should expect CellA_GFP and CellB_ME480 to show similar mechanical properties change in the defined RT-DC feature space, so for CellA_shM480 and CellB_ME480. In such a context, CellA_GFP and CellB_GFP are considered as

controls, because they are more similar to wild-type cell lines. Measurement was done 6 days after doxycycline treatment, and 15 minutes after the detachment from the adhesive states.

Table 4: Cell populations before and after filtering

Cell line (control*)	ME480	# before	# after
CellA_GFP*	endogenous	11882	8349
CellA_ShME480	silenced	3324	2811
CellB_GFP*	not exist	4871	4115
CellB_ME480	exogenous	3008	2580

As shown in **Figure 6(a)(b)**, cell populations were plotted across the size versus deformation plane (visualization inspired by (1)). CellA_GFP has a wider distribution, whereas the remaining three populations lie in specific regions within its regime. Compared to two ME480-absent conditions, cell lines with ME480 protein expression shifted toward a higher deformation distribution. Such a shift co-occurred with an area increase in cell A, which is not the case in cell B. The significance of size and deformation change was then confirmed by Welch's t-test in **Figure 6(c)(d)**.

Given the limited information, here we make a bold hypothesis: ME480 might be a component of the cytoskeleton or cytoskeleton-associated proteins that could directly alter cellular mechanics by rearranging the substructure of the cell body, making the cell easier to get deformed. Notably, *ME480* is not a crucial gene, because its absence does not cause lethality in both cell A and B.

Intensity metrics improves population classification

To further separate cell populations, we included intensity metrics: mean intensity and skewness (i.e., third order of spatial moment) into the embedded feature space. As shown in **Figure 6(e)(f)**, cell A showed significantly higher mean intensity than cell B, and the skewness is also heterogeneous. With the intensity profile, two kinds of visualizations are available. In high-dimensional space as in **Figure 6(g)**, cell A takes the high-intensity profile regime, while cell B enriches in the low-intensity domain. The same trend can also be highlighted in the tSNE map as in **Figure 6(h)**. Cell A and B align across the diagonal, in which cell type separation is more distinguishable than ME480-present/ME480-absent separation. Furthermore, two control groups are closer to each other compared to their experimental counterparts.

To assess the influence of each feature, we developed a linear model incorporating one interaction term for intensity metrics.

$$\begin{aligned}
 \text{Cell line} = & \beta_0 + \beta_1 \cdot \text{Area} + \beta_2 \cdot \text{Def} + \beta_3 \cdot \text{Mean} \\
 & + \beta_4 \cdot \text{Skew} + \beta_5 \cdot \text{Mean} \cdot \text{Skew}
 \end{aligned} \tag{1}$$

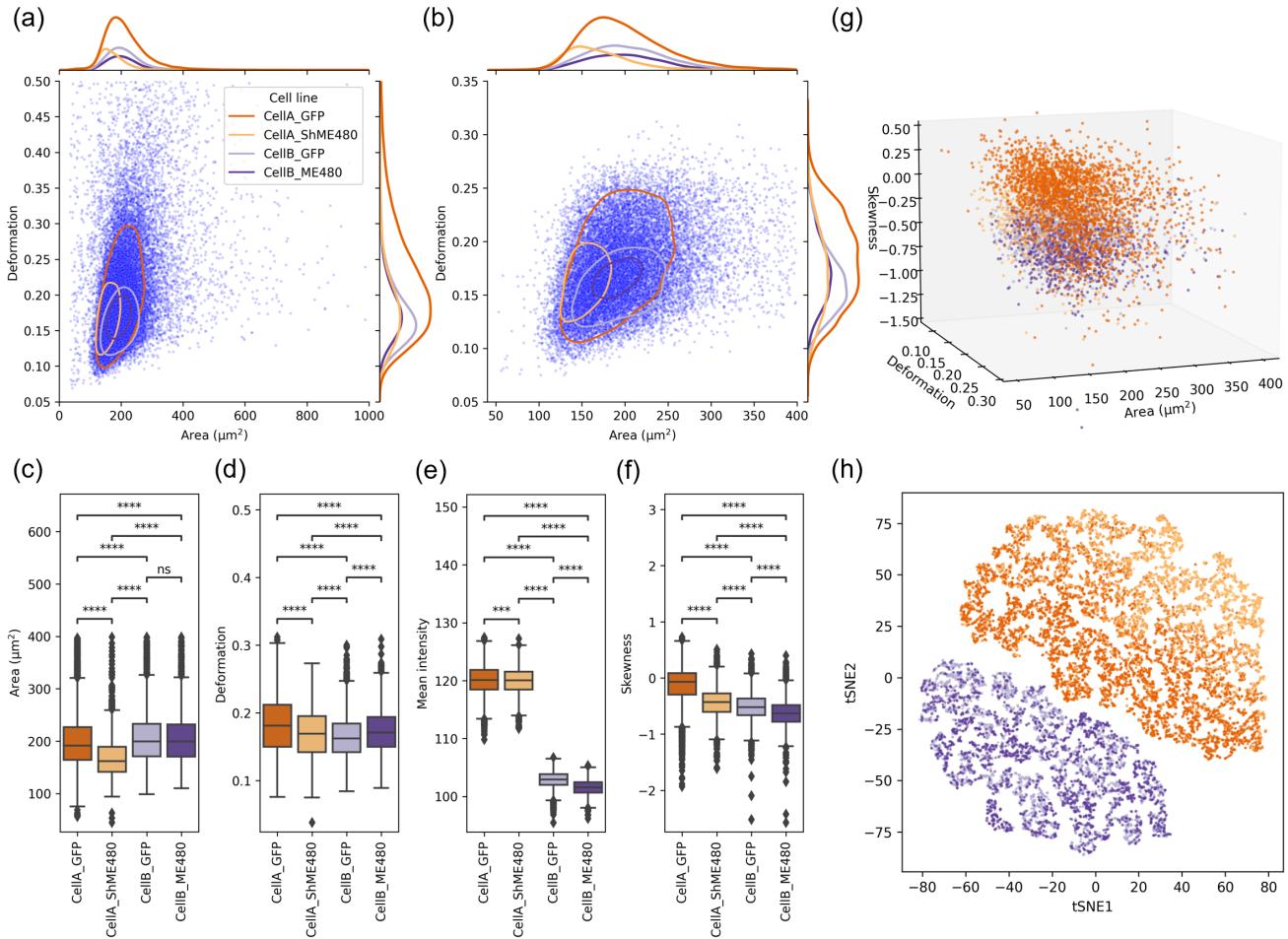


Figure 6: ME480 regulates higher deformation in a cell-type dependent manner. (a-b) Scatter plots of cross-sectional area vs. deformation for the unfiltered (a) and filtered (b) cell populations. All cells were plotted in the same color. 50%-density contours were highlighted for each cell line. Distributions in both axes are skewed with long tails toward the higher value range. The filtering step made the probability distributions of deformation in all populations more symmetric. (c-f) Boxplots of cross-sectional area (c), deformation (d), mean intensity (e), and skewness (f) for each cell line annotated with Welch's t-test significance (ns $p > 0.05$; *** $0.0001 < p \leq 0.001$; **** $p \leq 0.0001$). (g) 3D scatter plot with the additional skewness feature. For visualization purposes, 5000 randomly selected samples were shown. (h) tSNE map of the cell populations in the 4D embedded space.

The results, illustrated in **Table 5**, highlight mean intensity as the most significant predictor, followed by deformation, area, and, notably, the non-significant skewness.

Table 5: Linear model predicting cell lines

Metrics	Coefficient	P-value
Intercept	12.6444	0.00e+00
Area	0.0003	2.18e-05
Def	0.7248	4.01e-14
Skew	-0.2093	1.72e-01
Mean	-0.1060	0.00e+00
Skew:Mean	-0.0045	6.53e-04

In summary, the 4D feature space consisting of area, deformation, mean intensity, and skewness not only realizes the separation of cell types but also emphasizes the populational shifts derived from ME480 expression.

DISCUSSION

Decoding ME480 and other confounding factors

Despite the above discovery, we are aware of the defects of the analyses and exceptions of the hypothesis we made. Besides the observed distribution shift, it appears that in both cell lines, the experimental groups showed a narrower distribution than

the control counterparts. Such a phenomenon might be due to the cellular perturbation derived from genetic engineering. In experimental groups, the endogenous expression of ME480 proteins was directly altered, either silenced or overexpressed. In addition to the desired effect on specific cellular properties, perturbation might also deviate the cell population from other wild-type phenotypes that are not of interest. For example, overexpression of actin isoform proteins was found to lead to increased migration and invasion capacities (13). Removal of cytoskeleton-associated proteins from cell poles can facilitate faster cell cycle transition (14). These relevant studies reveal that perturbation of cytoskeleton dynamics can give rise to phenomena such as morphological change or cell cycle dysregulation. The DC setup is sensitive to such shifts (1). Therefore, there is a chance that ME480 does not directly impact cell mechanics; instead, it enriches the population in specific cellular states which are merely subsets of the control condition. Further measurements and functional studies need to be done.

Significance of statistical test

DC allows the measurement of thousands of cells within seconds. For such a large number of events, a conventional t-test might easily end up significant due to differences arising from small biological variations. Therefore it is crucial to ensure the reproducibility of the impact of given treatments by considering replicates and applying a linear mixed-effect model for statistical readouts (6). In our case, there is no replicate. We inspected this by artificially sampling each cell line into 3 replicates and constructing a linear mixed-effect model with replicates as all sources of variation (i.e., batch effects, noises...).

$$\text{Metric} = \beta_0 + \beta_1 \cdot \text{Cell line} + \beta_2 \cdot \text{Rep} \quad (2)$$

Metrics in **Equation (2)** can be any of the features in the embedded space. The results for area, deformation, mean intensity, and skewness were still significant ($p < 0.05$), supporting the statistics we proposed in previous sections.

Limitation of the analytical workflow

The current cellular identification and segmentation algorithm strongly depends on the presence of the bright halo derived from defocusing. As shown in **Figure 3**, the algorithm performed not so well when it comes to samples with incomplete halo, especially those with sharp cell poles. One possible alternative is to replace variance thresholding with contour tracing or circular Hough transform, with the tradeoff of higher computation complexity. In addition, the algorithm is also vulnerable to background shifts. If there is an obvious background intensity change or vibration of the microfluidic channels, the background subtraction would fail, leading to false detection of channel edges or irrelevant pixels. One pos-

sible solution is to iteratively update the background during the analysis.

CONCLUSION

Arguably the biggest appeal of mechanical phenotyping using label-free methods is to discover heterogenous, and especially less frequent cellular profiles in a mixed population. As demonstrated above, with prior knowledge of 4 different cell lines, our analytical workflow managed to extract and identify features that separate cell populations in high-dimensional space. We discovered that the expression of ME480 makes cells prone to be deformed, accompanied by cell-type dependent phenotypic changes. However, we also found confounding factors that could be detrimental to the method. To ensure the accuracy of DC measurements, various strategies are employed, including but not limited to: precise alignment of detachment times, steady image acquisition, incorporation of biological replicates, and the implementation of a more robust statistical model involving random variables. Furthermore, interestingly in our case, intensity profile became the most relevant feature to distinguish different cell lines. This discovery seems to go against the motivation of utilizing a deformation-based flow cytometry setup. Without knowing the ground truth label and the intensity profile of single-cell samples, it would be extremely hard to give deterministic predictions of the cell identity. To sum up, DC and RT-DC is still a powerful tool, but it should be done together with other conventional biological assays to deliver a holistic description of cellular phenotypes.

AUTHOR CONTRIBUTIONS

Yung-Cheng developed image processing algorithms and visualization. Clarisse and Antoine carried out statistical tests. Yung-Cheng, Clarisse, Nadim, Michel, and Antoine wrote the article.

SUPPLEMENTARY MATERIAL

All codes, tables, and figures are available on the [GitHub repository](#). Processed movies are accessible from the [Drive folder](#). Please contact [Clarisse](#) or [Yung-Cheng](#) if links are expired.

Table 6: Analytical workflow

Filename	Content
Part_A.ipynb	Image stack size reduction
Part_B.ipynb	Detachment time incertitude analysis
Part_C.ipynb	Segmentation, statistics, and visualization
Part_C_roi.ijm	ROI measurement

Table 7: Numerical and graphical outputs

Folder	Content
Data	All csv files generated and used for analysis
Figures	Png figures used in the article
Movies	Tif stacks for part A and C

REFERENCES

1. Otto, O., P. Rosendahl, A. Mietke, et al., 2015. Real-time deformability cytometry: on-the-fly cell mechanical phenotyping. *Nat Methods* 12:199–202.
2. Urbanska, M., P. Rosendahl, M. Kräter, and J. Guck, 2018. High-throughput single-cell mechanical phenotyping with real-time deformability cytometry. In *Methods in Cell Biology*, Elsevier, volume 147, 175–198.
3. Gossett, D. R., H. T. K. Tse, S. A. Lee, Y. Ying, A. G. Lindgren, O. O. Yang, J. Rao, A. T. Clark, and D. Di Carlo, 2012. Hydrodynamic stretching of single cells for large population mechanical phenotyping. *Proceedings of the National Academy of Sciences of the United States of America* 109:7630–7635.
4. Bradski, G., 2000. The OpenCV Library. *Dr. Dobb's Journal of Software Tools* .
5. Van der Walt, S., J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, 2014. scikit-image: image processing in Python. *PeerJ* 2:e453.
6. Hawley, T., and R. Hawley, editors, 2018. Flow Cytometry Protocols, volume 1678. Springer New York.
7. Buitinck, L., G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, 2013. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*. 108–122.
8. Charlier, F., M. Weber, D. Izak, E. Harkin, M. Magnus, J. Lalli, L. Fresnais, M. Chan, N. Markov, O. Am-salem, S. Proost, A. Krasoulis, getzze, and S. Repplinger, 2022. Statannotations. <https://doi.org/10.5281/zenodo.7213391>.
9. Seabold, S., and J. Perktold, 2010. statsmodels: Econometric and statistical modeling with python. In 9th Python in Science Conference.
10. Hunter, J. D., 2007. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* 9:90–95.
11. Waskom, M. L., 2021. seaborn: statistical data visualization. *Journal of Open Source Software* 6:3021. <https://doi.org/10.21105/joss.03021>.
12. Maloney, J. M., D. Nikova, F. Lautenschläger, E. Clarke, R. Langer, J. Guck, and K. J. V. Vliet, 2010. Mesenchymal Stem Cell Mechanics from the Attached to the Suspended State. *Biophysical Journal* 99:2479–2487.
13. Simiczyjew, A., A. J. Mazur, A. Popow-Woźniak, M. Malicka-Błaszkiewicz, and D. Nowak, 2014. Effect of overexpression of - and -actin isoforms on actin cytoskeleton organization and migration of human colon cancer cells. *Histochemistry and Cell Biology* 142:307–322.
14. McClatchey, A. I., 2014. ERM proteins at a glance. *Journal of Cell Science* 127:3199–3204.