A comprehensive suite of measures to quantify cell shape and actin organization

Running title: Shape and texture measures

Authors: Elaheh Alizadeh[1], Wenlong Xu[1], Jordan Castle[2], Jacqueline Foss[3,4], Ashok Prasad[1,3,**] .

1- Department of Chemical and Biological Engineering, 2- Department of Biology, 3-School of Biomedical Engineering, 4- Department of Mechanical engineering, Colorado State University, Fort Collins, CO, 80523.

**Corresponding author

KEY WORDS: Morphometrics, shape quantification, texture quantification, cell shape, Cytoskeleton, Cancer.

Summary statement

We develop a set of quantitative measures for studying changes in the shape and texture of cells, and experimentally validate their performance on osteosarcoma cells perturbed with different cytoskeletal drugs.

Abstract

A number of recent studies have shown that cell shape and cytoskeletal texture can be used as sensitive readouts of the physiological state of the cell. However, utilization of this information requires the development of quantitative measures that can describe relevant aspects of cell shape. In this paper we develop a set of quantitative measures, some used previously and others that are new, that have desirable properties for shape and texture quantification of cells. These measures, broadly classifiable into the categories of textural, spreading and irregularity measures, are tested by using them to discriminate between osteosarcoma cell lines treated with different cytoskeletal drugs. We find that even though specific classification tasks often rely on a few measures, these are not the same between all classification tasks, thus requiring the use of the entire suite of measures for classification and discrimination. We provide detailed descriptions of the measures, as well as codes to implement them. Image based quantitative analysis has the

potential to become a new field of biological data ("image-omics"), providing quantitative insight into cellular processes.

Introduction

The shape of a cell spread on a substrate is determined by the balance between the internal and external forces exerted on the cell boundary. The cell exerts forces and responds to external forces, from the extra-cellular matrix (ECM) or from neighboring cells, with the help of molecular motors and the cellular cytoskeleton, which is thus the ultimate determinant of cell shape (Fletcher & Mullins 2010a),(Pollard & Cooper 2009). The cytoskeleton is a complex network, made of three major kinds of filaments -- f-actin, microtubules and intermediate filaments -- that form a cross-linked dynamic meshwork in the cytoplasm, providing shape and structure to the cell (Fletcher & Mullins 2010b),(Ingber et al. 1994). The most dynamic constituent of the cytoskeleton, which is especially important in force generation and motility, is the filamentous actin (f-actin) network(Stricker et al. 2010). The f-actin network is directly involved in the formation of lamellipodia and filopodia through polymerization of f-actin against the cell membrane(Pollard & Borisy 2003). A third kind of cellular protrusions, blebs, are a result of the cortical actin network detaching from the cell membrane(Charras et al. 2008), and the convex shapes of adherent cells have been shown to result from myosin-II driven actin contractility(Bischofs et al. 2008). Cellular adhesion is maintained by an actin network that is often organized into contractile bundles called stress fibers(Pellegrin & Mellor 2007), which are classified into four types, dorsal and ventral fibers, transverse arcs and the perinuclear actin cap(Tojkander et al. 2012; Levy et al. 2010). Dorsal stress fibers are stabilizing structures associated with the cell edge, while the transverse arcs and the ventral fibers are major contractile structures, and the perinuclear actin cap regulates the shape of the nucleus(Tojkander et al. 2012).

The f-actin network is also deeply involved in force generation, force sensing and mechanotransduction. Contractile forces generated by myosin motors within cytoskeletal

networks, membrane extension caused by actin polymerization, changes in osmotic pressure by opening of water or ion channels are examples of internal forces that play a role in shape of a cell. External forces leading to shape changes are applied through neighboring cells or ECM (Paluch & Heisenberg 2009). Actin filaments can generate and also resist mechanical stresses and cell deformation. But they can also eventually reorganize and change their structure, thereby sometimes relaxing external stresses. Different mechanical properties of the cell cytoskeleton and ECM will lead to different shapes for the cell.

Thus the f-actin network is primarily responsible for the shape acquired by the adherent cell. It follows that the structure of the f-actin network must be related to the global shape of the cell, though the exact relation between the two is likely to be complex and non-linear.

Cell shape has been investigated as a marker and predictor of cellular phenotype and behavior. In our previous work, we used Zernike moments and geometric parameters as a measure of cell shape to distinguish between high metastatic and low metastatic osteosarcoma cancer cell lines with 99% accuracy(Alizadeh et al. 2016; Lyons et al. 2016). Other groups have also reported that cell shape can predict tumor grade (Sailem & Bakal 2017), changes in the nuclear/cytoplasmic ratio of NFkB (Sero et al. 2015), and YAP (Yes-associated protein), (Sero & Bakal 2017), chemosensitivity of human colon cancer cell lines(Pasqualato et al. 2013) ,(Pasqualato et al. 2012), differences in motility (Keren et al. 2008), forms of motility (Cooper et al. 2015), the progress of the Epithelial to Mesenchymal transition (EMT) (Ren et al. 2015) and the differentiation of Human Mesenchymal Stem Cells (hMSCs) into osteoblasts (Matsuoka et al. 2013).

Given the importance of the actin cytoskeleton in determining the shape of the cell and the nucleus however, the structure and organization of the f-actin network may provide additional information that can improve the prediction of cell physiology and better distinguish between cells in different states. There is a case to be made, therefore,

of including measures of actin organization into studies of cell shape and its relation with cell states, especially as actin staining is often used as the primary means to determine the shape of the cell. Actin staining also provides textural information, which is directly related with actin structure, and as we show below, addition of this textural information significantly improves the discrimination between cell types.

The importance of cell and nuclear morphology and cytoskeletal organization as a tool for understanding and predicting cell behavior raises the need to appropriately quantify the shape and cytoskeletal texture of a cell. Good quantitative measures, which we call cellular morphometrics, should be capable of capturing biologically important differences between different experimental conditions. However, not all measures will be optimal for each comparison, and thus one will have to begin with a larger set of quantitative measures and discard uninformative ones if needed. Furthermore, good measures must not show significant differences between technical replicates in the same experiment. Based on these arguments we chose to test and develop cellular morphometrics on cells treated with pharmacological modulators of the cytoskeleton. These quantitative measures also meet the criteria proposed for good morphometrics by Pincus et. al. (Pincus & Theriot 2007), i.e., that shape measures should possess fidelity, capture biologically important variation and be meaningful and interpretable.

Experimental Methods

For this study DUNN and DLM8 osteosarcoma cancer cells were used. The DLM8 line is derived from the DUNN cell line with selection for metastasis(Asai et al. 1998). Therefore, DLM8 is closely related to DUNN except for degree of its invasiveness. Both cell lines were a gift from Dr. Douglas Thamm (Colorado State University, CO, USA). They were cultured on glass detergent washed and air dried (GDA) substrates with standard culture conditions of 37°C and 5% carbon dioxide concentration in Dulbecco's Modified Eagle Medium (DMEM) (Sigma Alrdich). DMEM was supplemented with 10% EquaFETAL Fetal Bovine Serum (Atlas Biologicals) and 100 Units/ml penicillin

with 100 μg/ml streptomycin (Fisher Scientific-Hyclone). After 45 hours of culture, cells were incubated with different cytoskeletal drugs with description, conditions, and vendors listed in the **Table 1** for 3 hours. Then the cells were washed and fixed with 4% paraformaldehyde. Finally, they were fluorescently stained for nuclei (DAPI from Molecular probes) and actin (Acti-stain 488 phalloidin from Cytoskeleton, Inc). All the drugs were dissolved in Dimethyl sulfoxide (DMSO) (Fisher BioReagents™) and to drop its effect on cell shape and actin structure control study was also treated with DMSO with the same molarity as other drugs. Then the cells were imaged using fluorescent microscopy. Representative images of each cell line treated with these drugs are shown in **Figure 1**.

Image processing

In order to strike a balance between the throughput and the accuracy of the image processing, the image processing is fully supervised by the operator to reduce the number of artifacts and also automated as much as possible to speed-up the processing. The image processing code is stream-lined into a consecutive step-by-step workflow which consists of four steps as follows. 1) A graphic user interface (GUI) enables the binary thresholding of the actin and nucleus images. While a thresholding value is suggested automatically by Otsu's method, users can easily adjust the thresholding value using a slide bar in the GUI by visually checking the original image and the thresholded image displayed side-by-side in the GUI. 2) Cell declustering in the thresholded images is done using an optimized template of the open-source software CellProfiler (Kamentsky et al. 2011). 3) The outputs from the CellProfiler are then visually examined by the operator and corrections can be made if necessary using the modules functionalized into this step. To facilitate the following analyses, each cell is centered and saved separately into one 1024x1024 image. Except for CellProfiler, all the other image processing codes were programmed in-house using Matlab (Mathworks) and are available upon requests. Also

available are the detailed protocol on how to configure the image processing codes and the CellProfiler template used for cell declustering.

Data analysis methods

Quantifying changes in multidimensional shape space

Each cell can be represented as a point in multidimensional shape space. For ease of analysis this vector can be projected down to a lower-dimensional space of the first few principal components (PCs). Differences between different conditions or different cell lines can be represented by p-value between their shape distributions for each PC space. While doing multiple comparisons, we will often pick the most informative principal component for our analysis from among the first four PCs. This is done in the following manner. We pick the PC whose worst case, i.e. largest p-value among all comparisons, is better (i.e. smaller p-value) than that of any other PC, so that it is the best single measure for distinguishing between all the comparisons. In other words, for each principal component maximum p-value between all comparisons is calculated as Eqn 1.

$$\text{MaxP}_{(PC1)} = \text{Max}\ \{\text{p-value}_{(PC1,\,1)}\ ,\ \ldots,\ \text{p-value}_{(PC1,\,n)}\} \qquad (1)$$

Where $\text{p-value}_{(PC1,\,n)}$ is the p-value for n^{th} comparison in the first principal component. Then from the first four PCs, the one that has the smallest MaxP, maximum p-value between different comparisons, is chosen. In each shape category, we choose the best principal component by this criteria, and will describe them as the Primary Principal Components (PPC) of each shape quantification measure or category, for each analysis.

Pearson correlation

Pearson correlation coefficient is calculated between features as shown in Eqn 2. Here r is Pearson correlation coefficient, x_i is the feature x for i^{th} sample, y_i is the feature y for i^{th} sample, μ_x is the average value of the feature x for all of the samples, and μ_y is the average value of the feature y for all of the samples. This coefficient is calculated for all 14 cell lines and drug combinations (2 cell lines x 7 drugs) and the averaged coefficient is recorded. The results are shown in heat map plots in **Figure S.** 1. In these plots the

diagonal elements are correlation between features with themselves, so they are all 1. Also, Pearson correlation coefficient is symmetric so each heat map plot will be symmetric as well.

$$r = \frac{\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^{n}(x_i - \mu_x)^2}\sqrt{\sum_{i=1}^{n}(y_i - \mu_y)^2}} \quad (2)$$

Results

Developing Shape quantifiers

The actin structure is obtained from pixel intensities of labelled actin, which we can represent by a greyscale image (**Figure 2A**). These greyscale images of cells, or images of stained nuclei, are converted into a binary image (**Figure 2B**) and the coordinates of the edges are recorded (**Figure 2C**), yielding the cell boundary. To extract features of the morphology from this numerical data, we introduce three classes of shape measures, described as textural, spreading, and irregularity classes. Textural measures are divided into three subcategories called band based measurements, gray scale fractal dimension, and gray scale measures which are calculated based on the intensity plots of actin, **Figure 2A**. Spreading measures, which include measures based on a Zernike moment representation of shape as well as subcategories involving basic geometric parameters such as area and perimeter, are extracted from the binary representation of the cell(**Figure 2B**), its convex hull, or a similar image of the nuclei. Irregularity measures include parameters such as the waviness and roughness, which use information regarding the pixels at the boundary of the cell, **Figure 2C**. Each of these shape categories have been described below.

Textural measures

Band based measurements

This parameter is sensitive to changes in the radial symmetry of the distribution of actin filaments. We divide the image into 10 equally spaced concentric regions (Δr) around the center of mass of the image, called bands, as shown in **Figure S. 2**. Five

quantifiers are used to measure the differences in actin distribution between these bands. First, average intensity for each band is calculated. Then indices of the bands which have the lowest or the highest average intensity and their value are recorded. The last measurement is what is called above average adjusted intensity of the bands which is formulated in **Table S**. 1 along with other band based measures(J H Slater et al. 2015).

Band based measures are especially important in characterizing the actin distribution of the cells treated with Cytochalasin D, where actin has very unique symmetrical distribution. In these cells, dense foci are formed around the nucleus. The bands located in the central region of the cells are void of actin. At the outer bands, short linear actin structures aligned along the radius are observed.

Gray scale fractal dimension

Fractal dimension (FD) is a measure of roughness of objects, and can be applied to the characterization of texture in engineered and natural images. For planar images the fractal dimension lies between 2 and 3. There are many methods to calculate this measure, and we chose the box counting method for its inherent simplicity. In this method the binary image is first covered with an evenly spaced grid with side length of e. Then, the number of boxes which cover the fractal image are counted. This process is repeated by decreasing the side length and the FD is calculated based on Eqn 3

$$\mathrm{FD}=\lim_{e\to 0}\frac{\log N(e)}{\log\left(\frac{1}{e}\right)} \qquad (3)$$

Here e is the size of each box in the grid and the variable N (□) is the number of the boxes which contain the fractal.

We calculate FD based on a binarization of the greyscale image using edge detection methods to identify actin voids. This is done using four different edge detection methods in Matlab. To provide an example, binarized images of a cell from DUNN cell line treated with Cytochalasin D using these four edge detection methods are shown in **Figure S. 3**. We found that different edge detection methods pick up different aspects of

the actin structure, and so for every cell their gray scale image is binarized using these four methods, and their FD calculated (Costa 2013).

Other gray scale measures

Haralik et, al introduced a procedure for quantifying the texture of satellite images based on the spatial relation between the gray tone of neighboring pixels in an image. They used these measures for classifying the satellite images(Haralick et al. 1973). In this method the gray level co-occurrence matrix (GLCM, sometimes called Spatial-Dependence Matrix) is calculated for gray scale images. For an image that has intensities of 1, 2, …, g, the co-occurrence matrix is a g x g matrix such that its ij^{th} element is the number of the times that a pixel in the image has intensity equal to "i" and the pixel at a pre-defined distance $\vec{d}$ (which we choose to be 1 pixel) from it has intensity of j. An example of a 4 pixels x 4 pixels image with 5 gray tone levels is shown in **Figure S.** 4**.**

In a GLCM matrix when there are very few dominant transitions in gray tone of an image in neighboring pixels, the matrix will have small value for all of the entries. Each diagonal element of a GLCM matrix represents the number of the times that gray tone does not change in the neighboring pixel. Large diagonal elements imply that the image is homogenous. Large numbers in the far upper right and far lower left of the matrix implies large transitions in intensity and high contrast in an image. After calculating the GCLM matrix, shown for the example image in **Figure S.** 4**C**, 23 different measures are calculated to quantify the texture in an image. The list of measures are tabulated in **Table S. 2**, (Haralick et al. 1973), (Soh et al. 1999), (Uppuluri 2010). For example one of the parameters is the contrast. For interpretation of this measure, it is useful to keep in mind that the contrast in an image is proportional to the changes in gray tone, $n=|i-j|$, so the far upper right and far lower left which have bigger value of n will have bigger contribution to the contrast parameter. For homogenous images, the diagonal entries will be large, and will have a bigger contribution to the homogeneity measure.

In this study we calculate the GLCM matrix for the vector $\vec{d}$ equal to 1 pixel in magnitude and with directions of 0°, 45°, 90°, and 125° and then the average value for each parameter is reported. This process yields rotation invariant measures.

Spreading measures

In this class of shape quantifiers, the binary image of cell and nuclei is used. These measures include geometric parameters for nuclei and cell, convex hull measures and Zernike moments for cells.

Geometric measures

Parameters like area, perimeter, major axis of fitted ellipse, minor axis of fitted ellipse, and their ratio are examples of geometric measures. All geometric measures used in this study to quantify cell and nuclei's shapes are listed in **Table S4** and **Table S5**.

Convex Hull measurements

A convex polygon is a closed polygon such that the connecting line between any pair of points inside the polygon, lies completely inside the polygon. **Figure S. 5A** and **B** shows an example of a convex and non-convex polygon. The convex hull of a 2D shape is the smallest convex polygon that encloses the whole shape (Welzl & Gärtner 2014). An example of the convex hull of a 2D cell shape is shown in **Figure S. 5C**. Convex hull geometric measures used in this study are listed in **Table S6.**

Zernike moments

To calculate Zernike moments, the image of the cell is projected to the Zernike Polynomial basis function and the coefficients, called Zernike moments, are used as descriptor of cell shape. Zernike moments are complex numbers whose magnitude is rotation invariant, and can be made displacement invariant by moving the cell to the center of the imaging frame such that the center of mass of the image falls on the geometric center of the frame. The procedure has been discussed extensively previously (Alizadeh et al. 2016),and Zernike moments have been used previously in other types of biological image recognition (Tahmasbi et al. 2011).

Irregularity of boundary measures

In addition to cell spreading, irregularities in the boundary also carries information about the state of the cell. For example an irregular border could arise due to a large number of filopodia in the cell, which is a signature of a highly dynamic cytoskeleton. A highly contractile cell may retract from focal adhesions at the boundary creating many membrane protrusions that increase variability. To quantify irregularity of the cell boundary, the 2D boundary of the cell is used as shown in **Figure 2C**. We use the pixel positions that mark the boundary of the cell to calculate waviness, which estimates the periodic variation in the boundary, and roughness, which measures the non-periodic variation in the boundary, as discussed below.

Waviness measures

Using a Fourier series, a signal can be expanded in terms of linear combinations of orthogonal basis functions of sines and cosines with increasing frequencies. In the Fourier series expansion, it is assumed that the input signal is periodic. The outline of a cell is a closed curve, hence it is counted as periodic signal. The boundary of the cell could be represented in Cartesian coordinates, x and y, or polar coordinates, ρ and θ. Regardless of coordinate selection, it will lead to two independent signals which can be separately written as linear combination of cos and sin basis functions as Eqn 4.

$$f(x) = C_{x,0} + \sum_{n=1}^{nPixel} A_{x,n}\cos(wnx) + B_{y,n}\sin(wnx) \qquad (4)$$

Where nPixel is the number of the pixels in the boundary of a cell and $\frac{\mathbf{w}}{\mathbf{2\pi}}$ is fundamental frequency which is equal to $\frac{\mathbf{1}}{\mathbf{2\pi * nPixel}}$. The variable n is an integer number and n*w is the nth frequency in the decomposition. Variables $\mathbf{A_n}$ and $\mathbf{B_n}$ are amplitudes of $\mathbf{n^{th}}$ frequency, $C_{x,0}$ is the mean of the signal, and $\mathbf{f(x)}$ is the input signal.

If we use all the frequencies, we can reconstruct the cell. But the number of the descriptors will be too high. There is a tradeoff between the number of frequencies (shape

descriptors) used and the accuracy of the reconstruction. We are interested in quantifying shape without dealing with high number of parameters. Since the amplitude decreases with increasing frequency, we can filter higher frequencies and just use lower frequencies to decrease the number of descriptors used for the analysis. Here, with qualitative analysis of reconstructed cells with different frequency we decided to use only the first 35 frequencies as descriptor of the cell. Reconstruction of shape of a few cells using these 35 frequencies is excellent, as shown in **Figure 3.**

However, a remaining issue for the Fourier coefficients is that they are not rotationally invariant. We can construct a rotationally invariant measure by using Eqn 5, which removes the phase difference in the Fourier expansion.

$$\mathbf{C_n} = \sqrt{\mathbf{A_n^2} + \mathbf{B_n^2}} \qquad (5)$$

This also reduces the number of parameters by half. Even though reconstruction of cell shape will not be possible with the rotationally invariant measures, our results show that they are nevertheless useful parameters for distinguishing between cells with different degree of variations in the radius. We will refer to these parameters as Waviness parameters. Interestingly, for Fourier parameters, all the features show low correlation coefficient with each other (~0.35), suggesting that each coefficient carries mostly independent information. This observation was expected since we use orthogonal basis functions.

Roughness

As explained earlier, in the Fourier series decomposition of shape, higher frequencies have small amplitude and we neglect them. However, we can still obtain information about small amplitude variations from those high frequency terms. Following the method of Villanueva et. al., (Villanueva 2015) to account for high frequency measures, we reconstruct the cell with the 35 Fourier components and subtract it from the original signal. Now we can calculate the roughness measures from the remainder of the

signal. These parameters are also highly correlated with each other and are listed in **Table S. 7**.

Validating performance of developed measures on osteosarcoma cells perturbed with different cytoskeletal drugs.

Drugs which directly perturb microtubules, myosin-II or actin leave a unique signature on actin structure which is quantifiable by textural measures.

We found that the most dramatic and unique effects on cell shape and texture arise from drugs that directly perturb microtubules or actin filaments. As shown in **Fig. 1**, control cells acquire a polygonal shape on surfaces. Depolymerizing actin using Cytochalasin D leads to rounded cells where the actin has re-organized into alternating concentric rings of high and low density. Within each ring the actin is structured in very unique radial streaks! Increasing the depolymerization rate of microtubules using Nocodazole leads to a small rounded morphology. Inhibiting myosin II activity and decreasing cell's contractility using Blebbistatin leads to increase in irregularities in the cell boundary. These changes in shape and structure of a cell are similar for both DUNN and DLM8 cell lines. Changes in structure and morphology of the cells using other drugs are subtle and not easily identifiable by eye.

To quantify changes in shape and structure of the cells perturbed with different drugs, gray scale and binary image of the cells along with information of the boundary of the cells were used to measure all features in the 9 shape categories detailed above for each cell. **Table 2** demonstrates quantified changes in all the measures for all the drugs. As demonstrated in this table the changes in textural measures for all the drugs are significant. This means that all the drugs that perturbed actin either directly, or indirectly through perturbing other cytoskeletal components, lead to significant changes in actin cytoskeleton organization.

Actin reorganization changes the spreading measures for both the cell and the nucleus

To explore changes in 2D shape of a cell and nuclei by changing actin distribution we compare the PPC of the geometric measures of the cell and nuclei. As shown in **Table 2** perturbing actin significantly changes cell geometric measures other than DLM8 cell line treated with Jasplakinolide. Interestingly changes in actin structure not only changes cell geometric measures, but it also changes nuclei geometric measures for all conditions other than both cell lines treated with Blebbistatin and DUNN cell line treated with Jasplakinolide and PP2 drugs.

Zernike moments and Convex Hull parameters are similar in some respects. Although their quantification method is very different, as shown in reconstruction of the image of cells using Zernike moments (**Figure S. 6**) both measures ignore the irregularities and fine fluctuations in the boundary. As demonstrated in **Table 2** changes in the PPC for both measures are not distinguishable for DUNN cell lines treated with PP2, and both cell lines treated with FAKI 14. Moreover, hull geometric measures do not change significantly for DLM8 cells treated with Cytochalasin-D and DUNN cells treated with Jasplakinolide. In addition, Zernike moments do not change significantly for the DLM8 cells treated with Blebbistatin and DUNN cells treated with Cytochalasin-D. All other drugs lead to distinguishable changes in convex hull and Zernike moment measures.

Since many cell shape parameters are expected to be correlated with each other, we performed a correlation analysis of all the measures within each shape quantification category, which are shown in **Figure S.** 7. These results indicate that Zernike moments, cell geometric measures, nuclei geometric measures, and convex hull geometric measures are highly correlated with each other.

Cytoskeletal reorganization leads to changes in irregularities of a cell's boundary

As shown in **Table 2**, waviness measures for all the drug conditions, except for DUNN cell lines treated with Jasplakinolide or PP2 and both cell lines treated with FAKI 14, changes significantly with respect to controls. In addition, the roughness measures change significantly for both cell lines treated with Blebbistatin, Cytochalasin-D, FAKI 14, and Nocodazole. In both cell lines, Jasplakinolide does not change the roughness measure significantly, which is also the case for the DUNN cell lines treated with PP2.

Different categories of shape quantifiers represent partly non-redundant shape information

While we have shown that each of our major categories and subcategories serve as good shape measures, in that they can be used to look for interpretable shape changes in different drug conditions, it is not clear whether we need all of them to represent shape. In order to estimate the degree of redundant information carried by the different shape categories, we calculated the Pearson correlation coefficient between all the features from different shape categories (shown in **Table S8**). In general, features from two different shape categories are weakly correlated (<0.4) except for Convex Hull and Cell Geometric features which are highly correlated with each other. There are a few other specific exceptions for which the features are also highly correlated. They are as follows. Zernike moments with n<16 and m=0 are highly correlated with Area. The correlation coefficient between Cell Area and Zernike Moment (0_0) is 1. It decreases with increasing order till it is almost zero for Zernike Moment (22_0), then it becomes negative and increases in magnitude (**Figure S.** 7). The coefficient C_0 from waviness features has a correlation of 1 with the Mean Cell Radius. It is also highly correlated with Cell Area (0.93) and Convex Hull area. Since the Fourier decomposition is based on a radial representation of the cell shape, C_0 is a measure of the average cell radius, and should be expected to show these high correlations. Fractal Dimension measures also have high correlation with Cell Geometric, Nuclei Geometric, Gray Scale measures, and Zernike moments with m=0

and n<17. However, apart from these specific cases, the relatively high number of weak correlations between different shape categories implies that these shape categories contain non-redundant information about cell shape. Thus, quantitative shape analysis should ideally be carried out with representations from all of these shape categories in order for most efficient discrimination between different experimental conditions.

Discussion:

In this paper we introduce and provide the tools to quantify cell shape and cytoskeletal structure based on two dimensional images of cell morphology and actin structure. The Matlab toolbox used to process the images and quantify the shape and structure of a cell is shared in GitHub repository to be used by others (https://github.com/Elaheh-Alizadeh/Quantifiction-of-shape-and-structure), (https://github.com/Wenlong-Xu/Image_Processing_Cell_Shape). This work is among the few that quantifies structure of the actin. Using statistically significant differences between experimental conditions in principal component space as our measure, we explore the capacity of these quantitative measures in capturing biologically important information. We perturb the cytoskeleton of the cells with different drugs and explore their effect on cell shape and its structure. We first use textural measures to verify that actin structure of the cell changes in the cells treated by cytoskeletal drugs and then we explore changes in the cell's and nuclei's 2D shape and irregularities of the cell boundary accompanied by changes in actin structure. Here we showed that most of the drugs used in this study directly or indirectly lead to significant changes in actin structure. Then we explored effect of changes in structure of the cell on irregularities of cell boundary, nuclei spreading, and cell spreading. In most of the cases changes in actin structure are accompanied with significant changes in irregularities of cell boundary, and cell and nuclei spreading. The results show that textural measures and spreading measures are related but their relation is not simple, and the two classes of measures carry non-redundant information. It is worth mentioning that although we implemented textural

measures to quantify actin structure, it can be used to quantify other cell structures as well. These shape quantification methods presented in this paper will prove useful for computer aided diagnostics in diseases such as cancer that are associated with cytoskeletal perturbations, assessment of qualitative cellular changes in different experimental conditions, and for mechanistic understanding of the determination of cell shape. Textural measures introduced in this paper can also be used to quantify the texture of image of slices of tumors rather than single cells. The advantage of using these measures is that the results are more consistent, with a lower probability of human error and inconsistency. Furthermore these methods are fast and do not require expensive training.

Acknowledgement:

We gratefully acknowledge support from an NSF CAREER grant (PHY-1151454) to A. P.

List of Abbreviations:

Dimethyl Sulfoxide (DMSO)

Dulbecco's Modified Eagle Medium (DMEM)

Epithelial to Mesenchymal transition (EMT)

Extra-cellular Matrix (ECM)

Glass Detergent washed and Air dried (GDA)

Gray Level Co-occurrence Matrix (GLCM)

Human Mesenchymal Stem Cells (hMSCs)

Primary Principal Components (PPC)

Principal Components (PCs)

Figure 1: Representative images of DUNN and DLM8 osteosarcoma cancer cell lines with different drug treatments. Blue color represents nuclei and green color represents actin cytoskeleton.

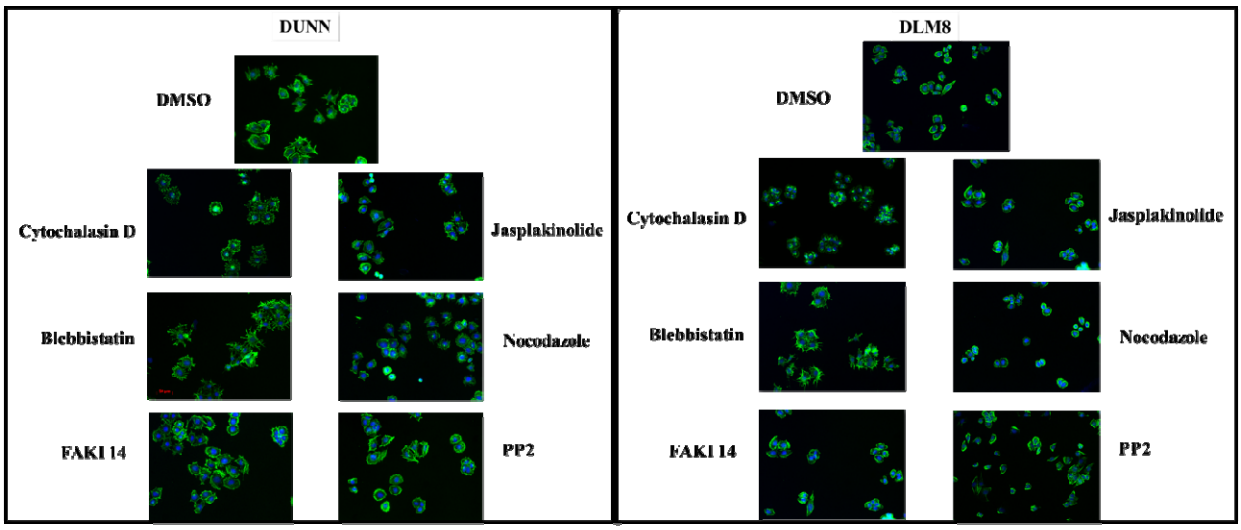


Figure 2: Representation of cell shape. A. Gray scale images using labeled actin. Intensity of each pixel is recorded and represented as a greyscale intensity plot. B. 2D outline: Position of each pixel in the boundary is recorded in polar or Cartesian coordinate.

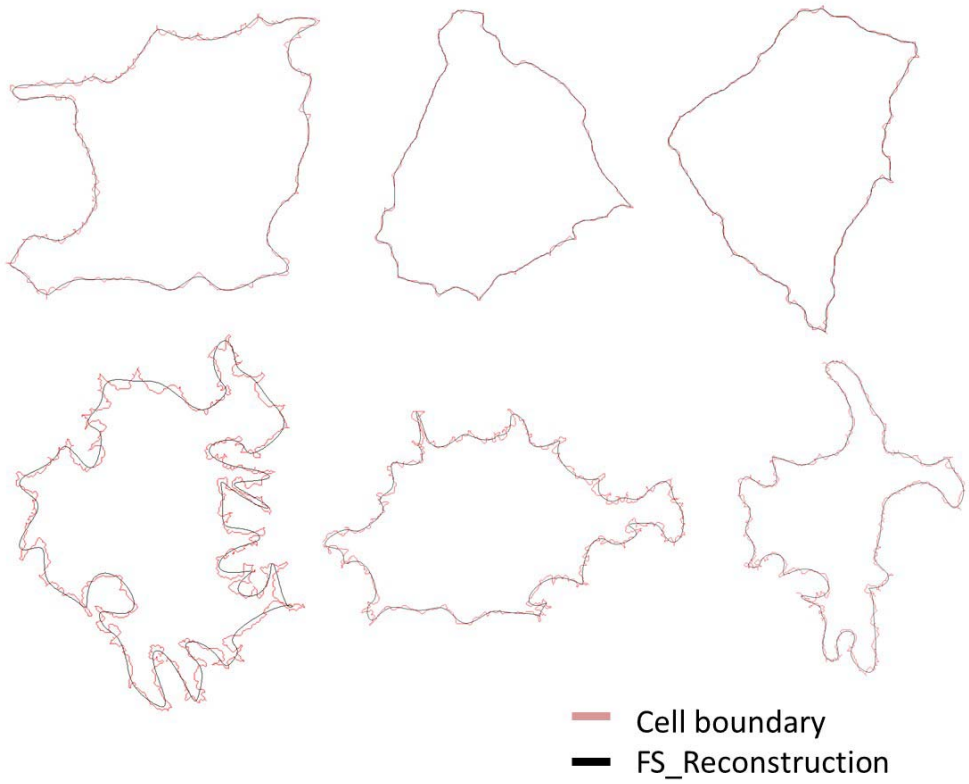


Figure 3: Reconstruction of the cells using Fourier decomposition. The red line is the actual boundary of the cell and black line is the reconstruction of the cell using the first 35 terms in the Fourier series expansion.

Table 1: Description of the drugs used to perturb cytoskeleton.

Drug	**Blebbistatin**	**Cytochalasin D**	**Jasplakinolide**	**FAKI 14**	**Nocodazole**	**PP2**
Target	Inhibits non-muscle myosin II activity	Inhibits actin polymerization	1. Promotes actin polymerization 2. stabilizes actin filaments	Inhibits focal adhesion kinase	Depolymerizing microtubules	Inhibits Src kinase
Application in cancer therapy	Inhibits cancer invasion(Nelson et al. 2011)	Used as chemotherapeutic agent[(Trendowski, Zoino, et al. 2015; Trendowski, Mitchell, et al. 2015).	Reduces lung metastases of systemic Lewis lung carcinoma(Takeuchi et al. 1998).	Showed anti-metastatic, anti-neoplastic and anti-angiogenic properties (Kang et al. 2015; Gabarra-Niecko et al. 2003; McLean et al. 2003; Hecker & Gladson 2003).	Has chemotherapeutic properties by arresting cells in mitosis(Endo et al. 2010).	Prevention of cancer metastasis(Nam et al. 2002).
Effect	1.Reduces F-actin contractility 2.Increase in cell	Decreases cell contractility	Stabilizes F-actin(Posey & Bierer 1999)	Reduces cell motility and focal adhesion	Cell shapes which lack tails and long protrusions	1. Induces strong cell-cell contact. 2.Enhances E-cadherin/catenin

	spreading and 3.Decreases cell motility(Sero et al. 2015)			turnover(Schlaepfer et al. 1999).	and decreased cell asymmetry(Rodionov et al. 1993).	expression which are strongly associated with the actin cytoskeleton(Nam et al. 2002).
Concentration (uM)	10	1	0.01	1	5	1
Vendor	Santa Cruz Biotechnology	Calbiochem	Calbiochem	Santa Cruz Biotechnology	Sigma Alrdich	Calbiochem

459

Table 2: Significance of changes in primary principal component (PPC) of morphometric classes for different drugs according to T-test analysis. The p-value of the null hypothesis that the experimental condition has the same mean as the control is reported for each class of morphometrics, based on the PPC. There are significant changes (p-value < 1E-3) in some actin textural measures (Band-Based, Fractal Dimension and Gray scale measures), spreading measures (Cell Geometric, Nuclei Geometric, Hull Geometric and Zernike moment) and Irregularity measures (Waviness and Roughness) for every experimental condition. Sample size is 300 cells for each cell line and drug condition other than DLM8 FAKI 14 which has a sample size of 200. In this table, P-values<1E-9 are highlighted in dark green, 1E-9 <P-values<1E-6 are highlighted in light green, and 1E-6 <P-values<1E-3 are highlighted in yellow.

Comparisons: Control vs	Cell Line	Band Based	Fractal Dimension	Gray Scale	Cell Geometric	Nuclei	Hull Geometric	Zernike Moment	Waviness	Roughness
Blebbistatin	DUNN	1.00E-07	2.60E-14	0.008	1.60E-78	0.006	3.10E-14	3.00E-21	2.00E-35	2.70E-45
	DLM8	3.00E-08	6.30E-45	6.00E-09	6.00E-25	0.004	7.90E-08	0.0018	2.00E-12	5.30E-16
Cytochalasin D	DUNN	3.00E-50	2.10E-31	2.00E-23	3.50E-21	2.00E-22	2.50E-06	0.0092	5.00E-10	3.30E-27
	DLM8	4.00E-31	2.10E-13	1.00E-23	4.50E-15	1.00E-18	0.05285	1.00E-05	3.00E-07	8.20E-07
FAKI 14	DUNN	2.00E-09	5.16E-31	1.00E-23	5.20E-09	2.00E-04	0.04184	0.1194	0.1398	2.20E-12
	DLM8	4.00E-04	6.60E-22	2.00E-21	1.50E-06	2.00E-06	0.07864	0.0997	0.3125	7.80E-07
Jasplakinolide	DUNN	2.00E-04	2.10E-13	9.00E-05	1.30E-08	0.001	0.02999	4.00E-05	0.0036	0.051
	DLM8	0.003	2.80E-09	2.00E-04	0.00207	5.00E-04	1.20E-05	0.0005	0.0001	0.02117
Nocodazole	DUNN	6.00E-13	5.60E-31	3.00E-25	1.30E-28	1.00E-04	4.60E-10	9.00E-07	1.00E-12	3.30E-16
	DLM8	0.001	9.80E-14	4.00E-07	1.20E-14	6.00E-06	3.00E-09	6.00E-14	3.00E-21	5.30E-13
PP2	DUNN	2.00E-21	1.40E-11	2.00E-10	2.50E-06	0.003	0.03602	0.0243	0.0089	0.00387
	DLM8	4.00E-07	0.00035	3.00E-09	0.00043	8.00E-10	2.70E-09	4.00E-13	4.00E-11	0.00078

Bibliography

Alizadeh, E. et al., 2016. Measuring systematic changes in invasive cancer cell shape using Zernike moments. *Integr. Biol.*, 8(11), pp.1183–1193. Available at: http://xlink.rsc.org/?DOI=C6IB00100A [Accessed January 8, 2017].

Asai, T. et al., 1998. Establishment and characterization of a murine osteosarcoma cell line (LM8) with high metastatic potential to the lung. *International Journal of Cancer*, 76(3), pp.418–422.

Bischofs, I.B. et al., 2008. Filamentous network mechanics and active contractility determine cell and tissue shape. *Biophysical journal*, 95(7), pp.3488–96. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18599642 [Accessed January 5, 2017].

Charras, G.T. et al., 2008. Life and times of a cellular bleb. *Biophysical Journal*, 94(5), pp.1836–1853.

Cooper, S. et al., 2015. Apolar and polar transitions drive the conversion between amoeboid and mesenchymal shapes in melanoma cells. *Molecular biology of the cell*, 26(22), pp.4163–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26310441 [Accessed February 14, 2017].

Costa, A., 2013. Hausdorff (Box-Counting) Fractal Dimension. Available at: https://www.mathworks.com/matlabcentral/fileexchange/30329-hausdorff--box-counting--fractal-dimension.

De, J. & Villanueva, 2015. DESIGN AND IMPLEMENTATION OF A CALCULATION PROGRAM IN MATLAB PLATFORM FOR EVALUATION OF ROUGHNESS. Available at: https://upcommons.upc.edu/bitstream/handle/2117/85844/Master’s Thesis JdV.pdf.

Endo, K. et al., 2010. Nocodazole induces mitotic cell death with apoptotic-like features in Saccharomyces cerevisiae. *FEBS Letters*, 584(11), pp.2387–2392. Available at:

http://www.sciencedirect.com/science/article/pii/S0014579310002991.

Fletcher, D.A. & Mullins, R.D., 2010a. Cell mechanics and the cytoskeleton. *Nature*, 463(7280), pp.485–92. Available at: http://www.nature.com/doifinder/10.1038/nature08908 [Accessed August 21, 2017].

Fletcher, D.A. & Mullins, R.D., 2010b. Cell mechanics and the cytoskeleton. *Nature*, 463(7280), pp.485–492.

Gabarra-Niecko, V., Schaller, M.D. & Dunty, J.M., 2003. FAK regulates biological processes important for the pathogenesis of cancer. *Cancer metastasis reviews*, 22(4), pp.359–74. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12884911 [Accessed June 22, 2017].

Haralick, R., SHANMUGAM, K. & ITS'HAK, D., 1973. Textural Features for Image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), pp.610–621.

Hecker, T.P. & Gladson, C.L., 2003. Focal adhesion kinase in cancer. *Frontiers in bioscience□: a journal and virtual library*, 8, pp.s705-14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12700131 [Accessed June 22, 2017].

Ingber, D.E., Sun, Z. & Betensky, H., 1994. E , Cytoskeletal ~ Ec ~ an ~ Cs ~ Control in Angiogene. *Science*, 7(12).

Kamentsky, L. et al., 2011. Improved structure, function and compatibility for cellprofiler: Modular high-throughput image analysis software. *Bioinformatics*, 27(8), pp.1179–1180.

Kang, C.G. et al., 2015. Rho-associated kinase signaling is required for osteopontin-induced cell invasion through inactivating cofilin in human non-small cell lung cancer cell lines. *Bioorganic & medicinal chemistry letters*, 25(9), pp.1956–60. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0960894X15002206

[Accessed June 22, 2017].

Keren, K. et al., 2008. Mechanism of shape determination in motile cells. *Nature*, 453(7194), pp.475–480.

Levy, J.R. et al., 2010. Cell adhesion: integrating cytoskeletal dynamics and cellular tension. *Molecular Cell*, 11(9), pp.633–643.

Lyons, S.M. et al., 2016. Changes in cell shape are correlated with metastatic potential in murine and human osteosarcomas. *Biology Open*.

Matlab, Properties of gray-level co-occurrence matrix. Available at: https://www.mathworks.com/help/images/ref/graycoprops.html?searchHighlight=glcm&s_tid=doc_srchtitle.

Matsuoka, F. et al., 2013. Morphology-Based Prediction of Osteogenic Differentiation Potential of Human Mesenchymal Stem Cells. *PLoS ONE*, 8(2).

McLean, G.W., Avizienyte, E. & Frame, M.C., 2003. Focal adhesion kinase as a potential target in oncology. *Expert opinion on pharmacotherapy*, 4(2), pp.227–34. Available at: http://www.tandfonline.com/doi/full/10.1517/14656566.4.2.227 [Accessed June 22, 2017].

Nam, J. et al., 2002. Src family kinase inhibitor PP2 restores the E-cadherin/catenin cell adhesion system in human cancer cells and reduces cancer metastasis. *Clinical Cancer Research*, 8(7), pp.2430–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12114449.

Nelson, T. et al., 2011. The myosin II inhibitor blebbistatin more potently inhibits 3-D cell invasion than 2-D cell migration in a human metastatic tumor cell. *Cancer Research*, 71(18 Supplement), pp.C1--C1. Available at: http://cancerres.aacrjournals.org/content/71/18_Supplement/C1.

Paluch, E. & Heisenberg, C.-P., 2009. Biology and Physics of Cell Shape Changes in Development. *Current Biology*, 19(17), pp.R790–R799. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0960982209014511 [Accessed August 21, 2017].

Pasqualato, A. et al., 2012. Quantitative shape analysis of chemoresistant colon cancer cells: Correlation between morphotype and phenotype. *Experimental Cell Research*, 318(7), pp.835–846. Available at: http://dx.doi.org/10.1016/j.yexcr.2012.01.022.

Pasqualato, A. et al., 2013. Shape in migration: Quantitative image analysis of migrating chemoresistant HCT-8 colon cancer cells. *Cell Adhesion and Migration*, 7(5), pp.450–459.

Pellegrin, S. & Mellor, H., 2007. Actin stress fibres. *Journal of Cell Science*, 120(20), pp.3491–3499. Available at: http://jcs.biologists.org/content/120/20/3491.abstract.

Pincus, Z. & Theriot, J.A., 2007. Comparison of quantitative methods for cell-shape analysis. *Journal of Microscopy*, 227(2), pp.140–156. Available at: http://doi.wiley.com/10.1111/j.1365-2818.2007.01799.x [Accessed March 19, 2017].

Pollard, T.D. & Borisy, G.G., 2003. Cellular Motility Driven by Assembly and Disassembly of Actin Filaments Cellular Motility Driven by Assembly and Disassembly of Actin Filaments. *Cell*, 112(4), pp.453–465.

Pollard, T.D. & Cooper, J.A., 2009. Actin, a Central Player in Cell Shape and Movement. *Science*, 326(5957). Available at: http://science.sciencemag.org/content/326/5957/1208.full [Accessed August 20, 2017].

Posey, S.C. & Bierer, B.E., 1999. Actin stabilization by jasplakinolide enhances apoptosis induced by cytokine deprivation. *The Journal of biological chemistry*, 274(7), pp.4259–4265.

Ren, Z.-X. et al., 2015. Suitable parameter choice on quantitative morphology of A549 cell in epithelial - mesenchymal transition. *Bioscience reports*, pp.1–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25899751.

Rodionov, V.I. et al., 1993. Microtubule-dependent Control of Cell Shape and Pseudopodial Activity Is Inhibited by the Antibody to Kinesin Motor Domain. , 123(6), pp.1811–1820. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2290860/pdf/jc12361811.pdf.

Sailem, H.Z. & Bakal, C., 2017. Identification of clinically predictive metagenes that encode components of a network coupling cell shape to transcription by image-omics. *Genome research*, 27(2), pp.196–207. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27864353 [Accessed February 14, 2017].

Schlaepfer, D.D., Hauck, C.R. & Sieg, D.J., 1999. Signaling through focal adhesion kinase. *Progress in biophysics and molecular biology*, 71(3–4), pp.435–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10354709 [Accessed June 22, 2017].

Sero, J.E. et al., 2015. Cell shape and the microenvironment regulate nuclear translocation of NF-κB in breast epithelial and tumor cells. *Molecular systems biology*, 11(3), p.790. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26148352 [Accessed February 28, 2017].

Sero, J.E. & Bakal, C., 2017. Multiparametric Analysis of Cell Shape Demonstrates that β-PIX Directly Couples YAP Activation to Extracellular Matrix Adhesion. *Cell Systems*, 4(1), p.84–96.e6. Available at: http://linkinghub.elsevier.com/retrieve/pii/S2405471216304100 [Accessed February 14, 2017].

Slater, J.H. et al., 2015. Recapitulation and Modulation of the Cellular Architecture of a User-Chosen Cell of Interest Using Cell-Derived, Biomimetic Patterning. *ACS Nano*,

9(6), pp.6128–6138. Available at: https://www.ncbi.nlm.nih.gov/pubmed/25988713.

Slater, J.H. et al., 2015. Recapitulation and Modulation of the Cellular Architecture of a User-Chosen Cell of Interest Using Cell-Derived, Biomimetic Patterning. *ACS Nano*, 9(6), pp.6128–6138. Available at: http://pubs.acs.org/doi/abs/10.1021/acsnano.5b01366 [Accessed February 23, 2017].

Soh, L., Tsatsoulis, C. & Member, S., 1999. Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices. *Ieee Transactions on Geoscience and Remote Sensing*, 37(2), pp.780–795. Available at: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=3CCA532A27525D5E9523557CD011F596?doi=10.1.1.224.3585&rep=rep1&type=pdf%5Cnhttp://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.224.3585.

Stricker, J., Falzone, T. & Gardel, M., 2010. Mechanics of the F-actin Cytoskeleton. *Journal of Biomechanics*, 43(1), pp.247–253.

Tahmasbi, A., Saki, F. & Shokouhi, S.B., 2011. Classification of benign and malignant masses based on Zernike moments. *Computers in Biology and Medicine*, 41(8), pp.726–735.

Takeuchi, H. et al., 1998. Jasplakinolide: interaction with radiation and hyperthermia in human prostate carcinoma and Lewis lung carcinoma. *Cancer chemotherapy and pharmacology*, 42(6), pp.491–496.

Tojkander, S., Gateva, G. & Lappalainen, P., 2012. Actin stress fibers - assembly, dynamics and biological roles. *Journal of Cell Science*, 125(8), pp.1855–1864. Available at: http://jcs.biologists.org/cgi/doi/10.1242/jcs.098087.

Trendowski, M., Mitchell, J.M., et al., 2015. Chemotherapy with cytochalasin congeners in vitro and in vivo against murine models. *Investigational New Drugs*, 33(2),

pp.290–299. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25563824 [Accessed June 22, 2017].

Trendowski, M., Zoino, J.N., et al., 2015. Preparation, In Vivo Administration, Dose-Limiting Toxicities, and Antineoplastic Activity of Cytochalasin B. *Translational Oncology*, 8(4), pp.308–317. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26310377 [Accessed June 22, 2017].

Uppuluri, A., 2010. GCLM feature calculation. Available at: https://www.mathworks.com/matlabcentral/fileexchange/22354-glcm-features4-m--vectorized-version-of-glcm-features1-m--with-code-changes-?s_tid=prof_contriblnk.

Welzl, E. & Gärtner, B., 2014. Computational Geometry. , pp.16–29. Available at: https://www.ti.inf.ethz.ch/ew/Lehre/CG13/index.html.

SUPPLEMENTARY INFORMATION

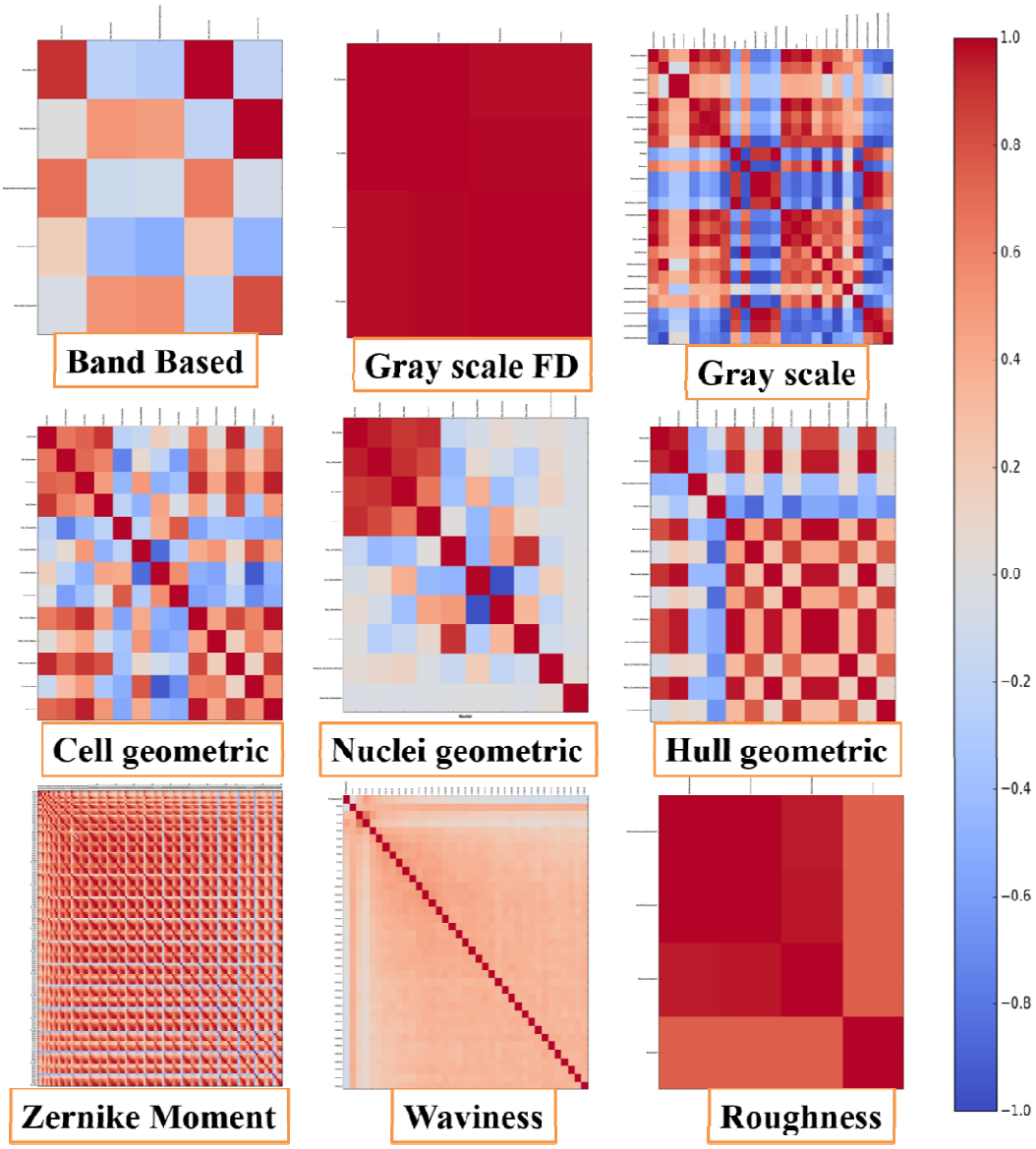


Figure S. 1 Pearson correlation coefficient within each shape category. Within each shape category features are highly correlated other than band based, waviness (Fourier), and some of the nuclei measures. Each pixel is the average of pearson correleation for 14 cases, 7drugs x 2 cell lines.

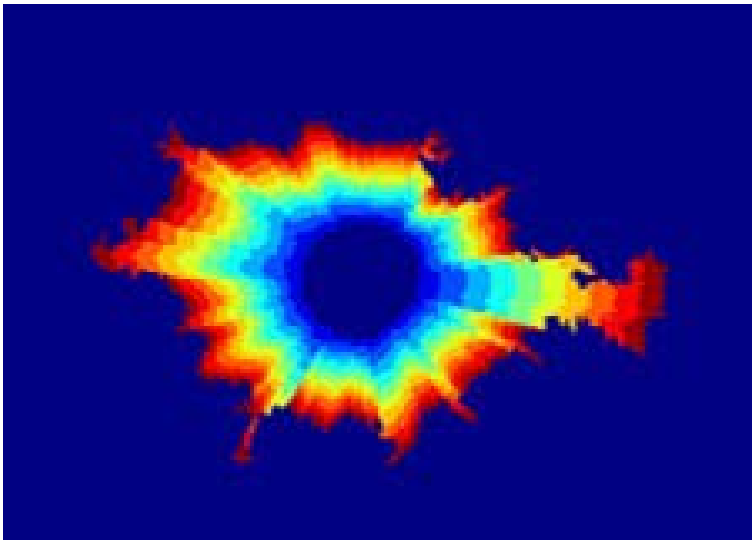


Figure S. 2: Band based division of textural image of a cell. Cell is divided into 10 equally spaced radial bands. All the measures calculated based on this representation is listed in **Table S**. 1.

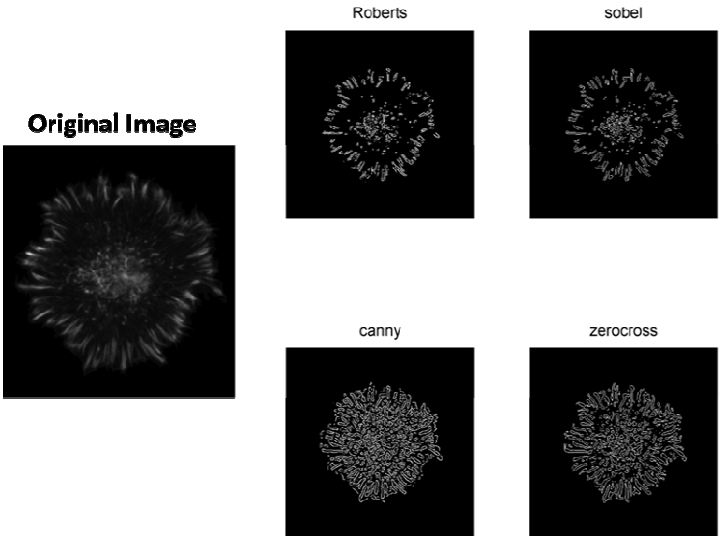
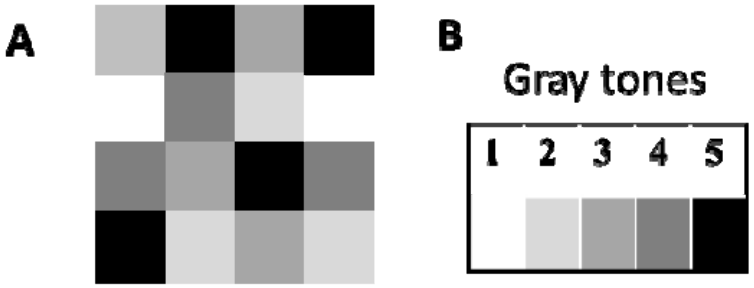


Figure S. 3: Matlab's Edging methods used to binerize gray scale image of actin. Four different methods of Roberts, Sobel, Canny, and zerocross are used in this paper.



2	5	3	5
1	4	2	1
4	3	5	4
5	2	3	2

D

Gray tone of : origin ↓ ,destination →	1	2	3	4	5
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)

Figure S. 4: Calculation of GCLM matrix. A) A cartoon of a 4pixels x 4pixels image. B) The image has 5 gray tones as shown. C) Numerical representation of image based on gray tone is shown. D) GCLM Matrix is 5 x 5 element and its ijth element is number of the times which pixel with intensity of j is in distance from the pixels with intensity of i. is an orbitrary vector but should be defined in advance. In this paper is difined to have length of one and it is defined in 4 directions of 0^0, 45^0, 90^0, and 135^0. Gray scale measures are calculated based on four directions and their average is reported.

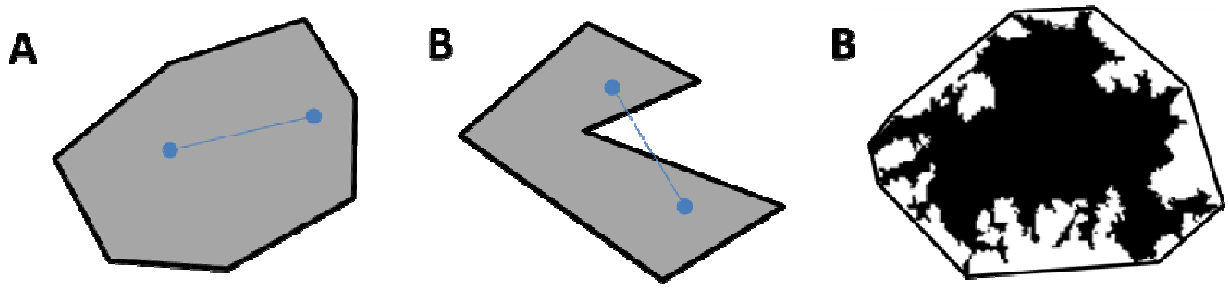


Figure S. 5: Convex Hull. A cartoon of a A) convex and B) non-convex polygon. A convex polygon is a polygon which any line connecting two points inside the polygon falls inside the polygon. C. Convex hull surrounding a cell.

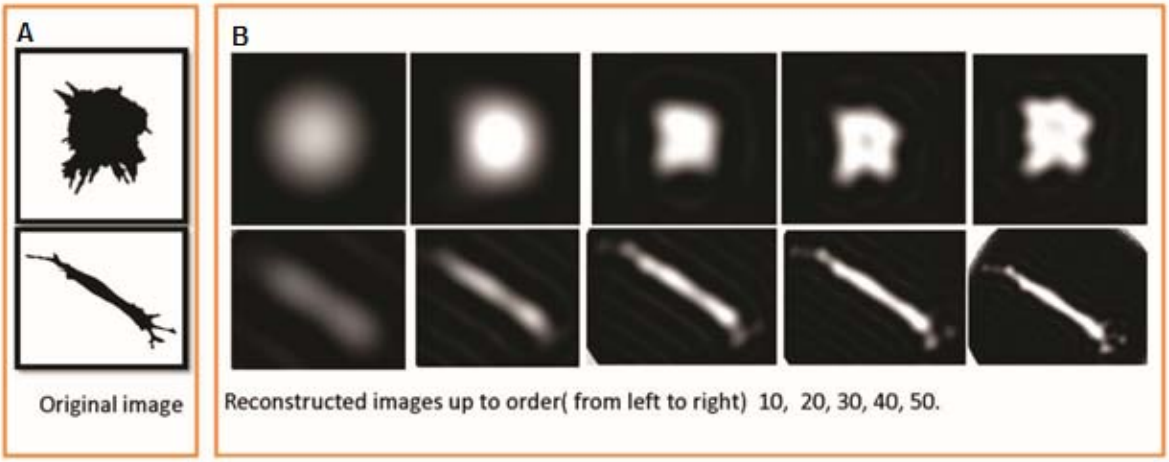


Figure S. 6: Reconstruction of cell images for two different shaped cells. A. Original images. B. Reconstructed images using Zernike moments up to order (from left to the right) 10, 20, 30, 40, 50.

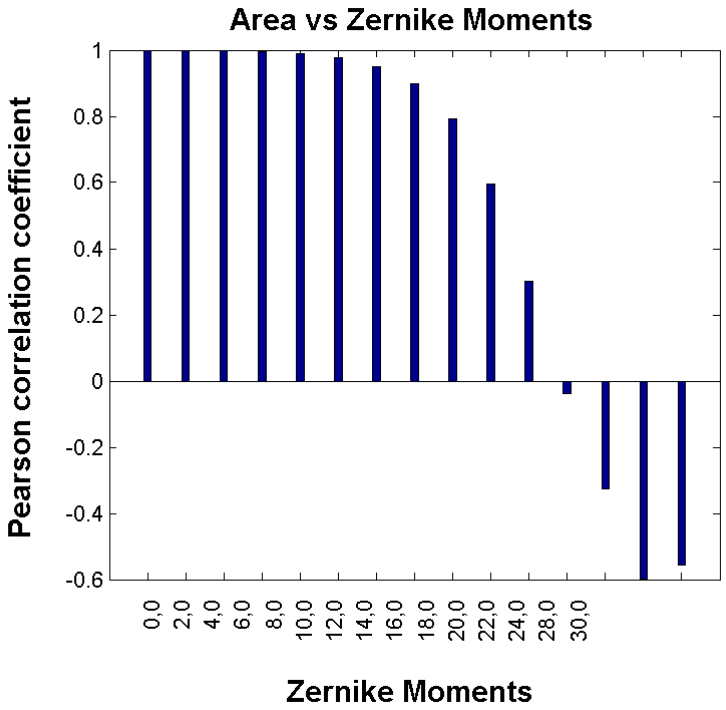


Figure S. 7: Pearson correlation coefficient between Zernike Moments with m=0 and Cell Area. Correlation is maximum and positive with Zernike moment of 0_0, it decreases with increasing the order and it is almost 0 for Zernike moment of 22_0 and it is negative for higher orders and increases in magnitude. Each data point is the average of pearson correleation for 14 cases, 7drugs x 2 cell lines.

Table S. 1: Band based measurements used in the paper(John H. Slater et al. 2015)**.**

Parameter		Formula
Max band intensity	Intensity of the band with maximum average intensity	Max $(\{I_{B_1}, \ldots, I_{B_{10}}\})$
Max band index	Location of the band with maximum average intensity	The n such that I_n = Max $(\{I_{B_1}, \ldots, I_{B_{10}}\})$
Minimum band intensity	Intensity of the band with minimum average intensity	Min $(\{I_{B_1}, \ldots, I_{B_{10}}\})$
Index of minimum band intensity	Location of the band with minimum average intensity	The m such that I_m = Min $(\{I_{B_1}, \ldots, I_{B_{10}}\})$
Above average adjusted intensity	Adjusted, weighted sum of average actin intensities greater than the mean intensity.	$\frac{\sum_{k=1}^{10} s(I_{B_k})\, I_{B_k}}{\text{Max}\left(\{I_{B_1}, \ldots, I_{B_{10}}\}\right) \sum_{k=1}^{10} s(I_{B_k})}$ Where $S(I_{B_k}) = \begin{cases} 1 & if\ I_{B_k} \geq mean(\{I_{B_1}, \ldots, I_{B_{10}}\}) \\ 0 & otherwise \end{cases}$

Table S. 2: Gray scale measures used in this paper (Matlab n.d.)**,** (Haralick et al. 1973)**,** (Soh et al. 1999)**,** (Uppuluri 2010)**.** See **Supplementary Table S. 3** for the parameters used in this table.

Parameter	Formula
Autocorrelation	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} ij\, p(i,j)$
Contrast_M(Matlab)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \lvert i-j\rvert^2\, p(i,j)$
Correlation_M(Matlab)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i\sigma_j}$
Correlation_P	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(ij)p(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y}$
Cluster Standard deviation	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i+j-\mu_x-\mu_y)^2 p(i,j)$
Cluster Prominence	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i+j-\mu_x-\mu_y)^4 p(i,j)$
Cluster Shade	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i+j-\mu_x-\mu_y)^3 p(i,j)$

Dissimilarity	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\{\lvert i-j\rvert p(i,j)\}$
Energy	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)^2$
Entropy	$-\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log\left(p(i,j)\right)$
Homogeneity_M(Matlab)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\frac{p(i,j)}{1+\lvert i-j\rvert}$
Homogeneity_P	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\frac{p(i,j)}{1+(i-j)^2}$
Maximum probability	$MAX_{i,j}\, p(i,j)$
Sum squares variance	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu)^2\, p(i,j)$
Sum_average	$\sum_{i=2}^{2N_g} i\, p_{x+y}(i)$
Sum variance	$\sum_{i=2}^{2N_g}(i - \mathrm{SumEntropy})^2\, p_{x+y}(i)$
Sum Entropy	$-\sum_{i=2}^{2N_g} p_{x+y}(i)\log\left(p_{x+y}(i)\right)$

Difference Variance	$\sum_{i=0}^{N_g-1} i^2 p_{x-y}(i)$
Difference Entropy	$-\sum_{i=0}^{N_g-1} p_{x-y}(i) \log\left(p_{x-y}(i)\right)$
Information Measure of Correlation1	$\frac{\text{Entropy} - HXY1}{\max\{HX, HY\}}$
Information measure Correlation 2	$\sqrt{(1 - e^{-2(HXY2-HXY)})}$
Inverse Difference Normalized(INN) (Matlab)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{p(i,j)}{1+\frac{\lvert i-j\rvert}{N_g}}$
Inverse Difference Moment Normalized(Matlab)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{p(i,j)}{1+\frac{(i-j)^2}{N_g^2}}$

Table S. 3: Parameters used in the definitions of gray scale measures in the Supplementary Table S. 2.

Parameter	Formula
P(i,j)	ijth entry of co-occurance probability matrix
N_g	Highest intensity of gray level
μ	Mean value of p(i,j)
$p_x(i)$	$\sum_{j=1}^{N_g} p(i,j)$
$p_y(i)$	$\sum_{i=1}^{N_g} p(i,j)$
μ_x	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} i\, p(i,j)$
μ_y	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} j\, p(i,j)$
σ_x^2	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} (i-\mu_x)^2\, p(i,j)$
σ_y^2	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} (i-\mu_y)^2\, p(i,j)$
$p_{x+y}(k)$	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)$

	Where i+j=k, k=2,..,2N_g
$p_{x-y}(k)$	$$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)$$ Where i-j=k, k=0,.., N_g-1
HX	$$-\sum_{i=1}^{N_g} p_x(i)\log(p_x(i))$$
HY	$$-\sum_{i=1}^{N_g} p_y(i)\log(p_y(i))$$
HXY1	$$-\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log(p_x(i)p_y(j))$$
HXY2	$$-\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p_x(i)p_y(j)\log(p_x(i)p_y(j))$$

Table S. 4: Geometric parameters.

Parameter	Description	Unit
Cell area	Area of the cell	Pixel^2
Cell perimeter	Perimeter of the cell	Pixel
Cell major axis	The major axis of an ellipse drawn around the cell	Pixel
Cell minor axis	The minor axis of an ellipse drawn around the cell	Pixel
Cell circularity (Smoothness)	$\frac{4\pi\ Area}{Perimeter^2}$	Unit-less
Cell aspect Ratio	$\frac{\text{Cell Major axis}}{\text{Cell Minor axis}}$	Unit-less
Cell roundness	$\frac{4\ Area}{\pi\ \text{Cell major axis}^{\ 2}}$	Unit-less
Cell solidity	$\frac{Convex\ hull\ area}{Cell\ area}$	Unit-less
Max cell Radius	Maximum radius of the cell	Pixel
Cell radius ratio	$\frac{\text{Max radius of the cell}}{\text{Min radius of the cell}}$	Unit-less
Mean Cell Radius	The mean of all radii drawn from the hull's centroid to an exterior point	Pixel
CV Cell Radius	The relative variation of radii drawn from the cell's center to an exterior point.	Unit-less

	Given by the Standard of Deviation of all Radii divided by the mean of all radii	
Max Span	The maximum distance from one point on the convex hull to another	Pixel

Table S. 5: Nuclei geometric parameters.

Parameter	**Description**	**Unit**
Nuc area	Area of the nucleus	Pixel^2
Nuc perimeter	Perimeter of the nucleus	Pixel
Nuc major	Major axis of an ellipse drawn around the nucleus	Pixel
Nuc minor	Minor axis of an ellipse drawn around the nucleus	Pixel
Nuc circularity	$\frac{4\pi\ Nuc\ \ area}{Nuc\ perimeter^2}$	Unit-less
Nuc aspect ratio	$\frac{\text{Nuc major}}{\text{Nuc minor}}$	Unit-less
Nuc roundness	$\frac{4\ \text{Nuc area}}{\pi\ \text{Nuc major}^2}$	Unit-less
Nuc solidity	$\frac{Area\ of\ convex\ hull\ of\ nuclei}{\text{Nuc area}}$	Unit-less
Distance nuc to cell centroid	Distance between the centroids of cell and nucleus	Pixel
Nuc to cell orientation	Angle between the distance nuc to cell centroid and cell major axis	Degree

Table S. 6: Convex hull geometric parameters.

Parameter	Description	Unit
Hull area	Area of the convex hull	Pixel^2
Hull perimeter	Perimeter of the convex hull.	Pixel
Ratio of perimeter of Hull to cell	Ratio of hull perimeter to cell perimeter.	Unit-less
Hull circularity	$\frac{4\pi\ Hull\ area}{Hull\ perimeter^2}$	Unit-less
Max hull Radius	Maximum distance from centroid of Hull to an exterior point on the hull.	Pixel
Hull radius ratio	Maximum/minimum radius of hull.	Unit-less
Mean hull radius	The mean of all radii drawn from the hull's centroid to an exterior point.	Pixel
CV hull radius	$\frac{STD\ hull\ radius}{Mean\ hull\ radius}$	Unit-less
Bounding circle diameter	The diameter of the bounding circle drawn around the cell.	Pixel
Max circle to hull radius	The maximum distance from the center of the bounding circle to an edge of the convex hull.	Pixel

Circle to hull radius ratio	$\frac{\text{Max circle to hull radius}}{\text{Min circle to hull radius}}$	Unit-less
Mean circle to hull radius	The mean of all radii drawn from the circle's center to the hull	Pixel
CV_Circle2Hull_Radius	$\frac{STD(\text{ all radii drawn from the circle's center to the hull })}{\text{Mean circle to hull radius}}$	Unit-less

Table S. 7: Roughness measurements.

Parameter	Formula	Units
Mean of absolute values	$\mathrm{Ra} = \frac{1}{n}\sum_{i=1}^{n} \lvert y_i \rvert$	Pixel
Root Mean Squared	$\mathrm{Ra} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2}$	Pixel
Maximum height of the profile	$R_{yt} = R_p - R_v$	Pixel
Kurtosis	$R_{Ku} = \frac{1}{nR_q^4}\sum_{i=1}^{n} y_i^4$	Pixel^4