

MORPHEUS:

AN AUTOMATED TOOL FOR UNBIASED AND REPRODUCIBLE CELL MORPHOMETRY

Supplementary Materials

Federico Alessandro Ruffinatti^{1*}, Tullio Genova^{2,3*}, Federico Mussano² & Luca Munaron^{3§}

¹ Department of Pharmaceutical Sciences, University of Eastern Piedmont, Novara, Italy; ² CIR Dental School, Department of Surgical Sciences, University of Turin; ³ Department of Life Sciences and Systems Biology, University of Turin; *These authors equally contributed; § Corresponding author.

Table of Contents

1.	Algorithm Details	3
1.1.	Launching MORPHEUS.....	3
1.2.	Segmentation and SNR	4
1.3.	Morphometry Analysis	8
1.4.	Orientation Analysis.....	11
1.5.	MORPHEUS Log.....	15
2.	Statistical Validation	17
3.	Experimental Validation	23
3.1.	Cell Cultures and Image Acquisition	23
3.2.	Morphometry	23
3.3.	Tolerance Effect.....	34
3.4.	Orientation Analysis.....	35
4.	Angular Resolution and Scale Dependency	42
5.	References	48

1. Algorithm Details

1.1. Launching MORPHEUS

MORPHEUS plugin can be easily integrated into Fiji menu, simply by copying *Morpheus_.ijm* file into the Fiji's *plugins* subfolder. Once installed, MORPHEUS can be launched from the *Plugins* menu and a dialog window will appear (Figure 1a, main text). Through this window, it is possible to specify the following parameters:

- the path of the input directory—the folder containing the images to be analyzed (hereafter denoted as *<input>*);
- the path of the output directory—the folder in which the results of the analysis will be saved (hereafter denoted as *<output>*);
- image file suffix—the extension of the images to be analyzed; *.tif* is the default option (this field and the related search are case insensitive);
- nucleus file identifier—a string of characters (where *dapi* is the default) that must be present in the file name of all the images stained for nucleus, but NOT in the images stained for cytoskeleton, allowing MORPHEUS to distinguish among them (also this field and the related search are case insensitive);
- the anti-spot lower bound ε —the minimum number of pixels a segmented object must have to be not discarded as ‘spot noise’ ($\varepsilon = 200 \text{ pixel}^2$ is the default value)¹;

¹ This value was obtained empirically, having proven to be effective with most of the datasets we had to deal with, however it cannot be considered an always-working value. Since this area is used as a first filter to get rid of the many non-cellular spots arising from segmentation (cell fragments, extracellular matrix debris, etc.), the user should preliminarily ‘guess’ the radius of the smallest cell that is likely to be found in that particular experiment and then set a value r just below it. In our case (see Section 3, *Experimental Validation*), we chose $r = 3.6 \mu\text{m}$ and worked with a microscopy featuring a conversion factor of $c = 2.2026 \frac{\text{pixel}}{\mu\text{m}}$. Hence, $\varepsilon = \pi(r c)^2 \approx 200 \text{ pixel}^2$. Clearly, the larger the cells or the higher the resolution, the larger ε could be. Even if conceptually straightforward, such a computation is only possible when the factor c is known, but often this is not the case. Since c is the result of many magnification stages (objective, downstream optical systems, digital camera, and so on), typically it can be reliably estimated just by direct calibration. However, in the everyday laboratory practice, the experimenter could be unaware of—or not interested in—such a value. For this reason, MORPHEUS uses pixels as ‘neutral’ unit of measure

- tolerance value T —the number of standard deviations from the mean of the sampling distributions used by MORPHEUS to define the characteristic range of area and circularity for the ‘typical’ isolated cell. In other words, it determines the confidence level for the estimates of the parameters of interest (namely \bar{M}_A , $\overline{\min}_C$, and $\overline{\max}_C$, according to the *Unsupervised Learning Step* described in the main text). T can take discrete values from 1 to 6: the larger the value of T , the more tolerant the selection process will be²;
- 4 checkboxes that allow the user to choose which optional MORPHEUS functions to enable among Orientation analysis, Nucleus analysis, Full output and Verbose log.

1.2. Segmentation and SNR

When clicking OK button, MORPHEUS starts scanning the user-defined input folder searching for the images stained for cytoskeleton and no further intervention by the user is required until the analysis of the whole dataset is completed.

As a first step, a separation of the objects of interest from the background—namely a segmentation process—is performed. Even if there are no general methods, the segmentation implemented in MORPHEUS represents a quite standard protocol in image analysis [1]. Typically, prior to thresholding, a

and we strongly suggest adjusting ε value empirically, knowing that the default value is a good starting point and an effective choice in most of the practical situations.

² Assuming normality for the **means** of sample median area and circularity extreme values (which is justified by the central limit theorem), it turns out that the two-sided critical region lying within $T = \{1, 2, 3, 4, 5, 6\}$ standard deviations from the mean encompasses, respectively, the $\{68.27, 95.45, 99.73, 99.993, 99.99994, 99.999998\}$ percentage of the whole distribution. Since a 99.999998% confidence interval contains the true parameter value nearly all the time (the outside-range occurrence rate is 1/506,797,346, that is to say one day every 1.38 million years or two days since the appearance of *Homo Habilis!*), the value $\tau = 6$ was chosen as the highest confidence level allowed by MORPHEUS. Moreover, the integer nature of T parameter is purely conventional, but it was established to limit and simplify the user choice and increase reproducibility.

preprocessing step aiming at increasing the separability between object and background pixels is required to get better results. This is achieved either by increasing the difference between the two classes (e.g. by contrast enhancement) or by increasing the homogeneity within each class (e.g. by smoothing). In this regard, noise and non-uniform illumination are the two major sources of heterogeneity in input images, possibly hindering the automated detection of the correct intensity threshold during segmentation and eventually leading to background/foreground misclassification. Specifically, MORPHEUS implements segmentation as a procedure in five successive steps (Figure 1b, main text):

1. **Contrast enhancement.** This sub-step is carried out using the *Auto* option provided by Fiji/ImageJ through the *Brightness/Contrast...* panel (from *Image* → *Adjust* menu). Thus, brightness and contrast are automatically adjusted based on the image's histogram and intensity values are remapped over the entire available domain (256 and 65536 gray levels for 8 and 16-bit images, respectively). The optimization is done by allowing a small percentage of pixels (0.35%) to become saturated (black or white), in order to get rid of possible intensity outliers and exploit the full dynamic range.
2. **Background subtraction.** MORPHEUS accomplishes such a task using Sternberg's *rolling ball* algorithm [2] in order to compensate for uneven illuminated background (i.e. removal of low-frequency intensity fluctuations).
3. **Smoothing.** Noise attenuation is performed through a standard 3×3 mean convolution filter (i.e. removal of high-frequency intensity fluctuations).
4. **Thresholding.** Since using the same threshold over a set of images is never recommended (because of the inherent inter-image variability) an image-specific threshold value should be computed for each image by an automated selection technique aimed at optimizing some objective criterion. In particular, MORPHEUS determines the threshold value by Li's method for the minimization of cross-entropy [3].

5. **Binarization.** The image is converted to black and white (black objects on white background) to make it suitable for *Analyze Particle* function (see below *Morphometry Analysis* section).

The result of such a segmentation process is saved as an image into the *Cell* subfolder of the output directory, using the conventional name notation *Img_#a - Segmentation - <filename>.tif*, where $\# = 1, 2, \dots, m$ is a progressive integer identifying the processed image and *<filename>* is its original name.

Segmentation, and thresholding sub-step in particular, is surely one of the most critical steps in MORPHEUS pipeline, since its possible failure is likely to heavily impinge on the downstream analysis. In our experience, Li's algorithm seemed to give the best performance, especially when associated with the other steps of MORPHEUS segmentation. However, other auto-thresholding algorithms provided by Fiji, such as Huang's, Renyi Entropy, Yen's and Triangle method, led to very similar results. Overall, this particular segmentation pipeline has been assembled to be general enough for being effective in most of the practical situations and maintaining at the same time the maximum level of automatization, since no parameter needs to be specified by the user. In any case, after MORPHEUS terminates, it is always recommended to check the results of both segmentation and cell detection. Badly thresholded images can be easily identified by inspecting the output folder content and their possible presence usually indicates an insufficient image quality or, more formally, a too low Signal-to-Noise Ratio (SNR), as stated in point 2 of *Sample Preparation* section in the main text.

Actually, there is no universal definition for the calculation of SNR of an image. In fact, without a ‘truly’ noiseless counter-image used as reference or any a-priori-knowledge about the structures of interest, SNR computation can be very hard—even if not impossible—to be accomplished. In many real cases, only rough estimations can be obtained. Nevertheless, in the case of biomedical images featuring many zero-intensity background pixels, a good SNR index can be defined as the ratio of the mean value intensity of a reference structure (e.g. a cell) to the

background standard deviation. More precisely, defining the two random variables:

B = pixel intensity in the background region;

R = pixel intensity in the reference region;

SNR can be defined as:

$$\text{SNR} = \frac{\bar{R} - \bar{B}}{\sigma_B}$$

where σ_B is the standard deviation of B (i.e. the noise).

This value can be easily estimated by the user through basic Fiji/ImageJ functions. Just use the *Selection tool* to manually select some cells and a background region. Then get the statistics of interest for R and B , respectively, through *Measure* command from the *Analyze menu* (Figure SM1). Using such a procedure on our validation datasets (see Section 3 below), we estimated that a value of at least SNR = 10 is needed to get a good segmentation.

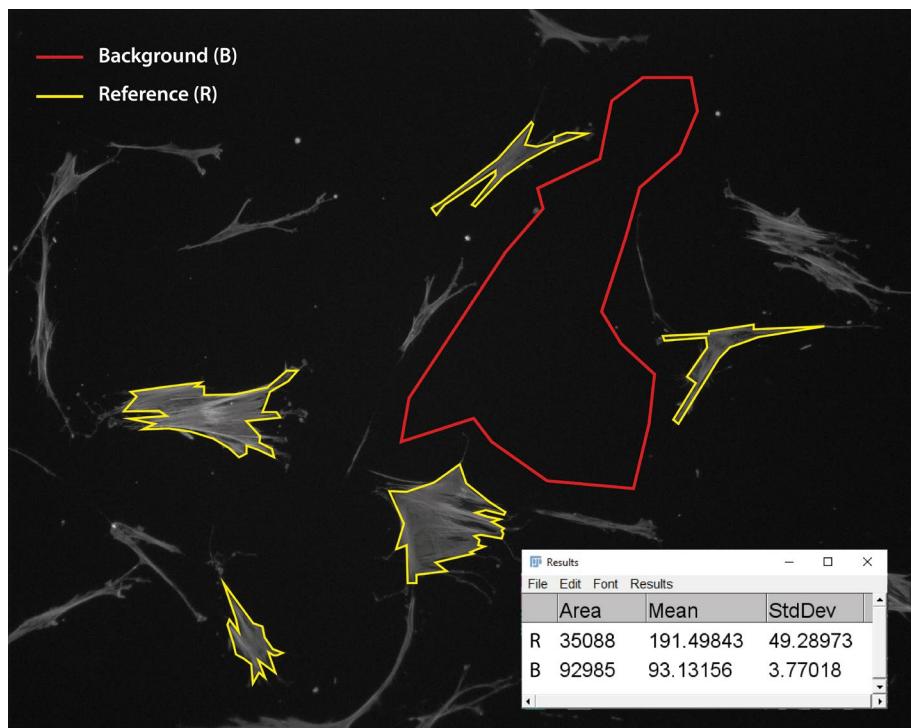


Figure SM1. Example of SNR manual estimation in Fiji/ImageJ environment. First, cells chosen as reference signal are selected (R , yellow lines) and mean intensity is computed by Measure command. Then mean and standard deviation of pixel intensity is computed for a background region (B , red line). Finally, SNR can be computed according to the definition given above. For this exemplificative image is $\text{SNR} = (191.5 - 93.1)/3.8 = 26.1$.

1.3. Morphometry Analysis

After the unsupervised learning step (see *Algorithm Flow* section in the main text), MORPHEUS should be able to recognize isolated cells among all the objects present in each sample image. As already stated, this is done by a second run of *Analyze Particle* function over the whole dataset, using $[A_{low}, A_{high}]$ and $[C_{low}, C_{high}]$ ranges as filtering parameters in input (Figure 1f, main text). The results of such a selection are saved as images into the *Cell* subfolder of the output directory, using the conventional name notation *Img_#b - Detection - <filename>.tif*, where # is the progressive ID and *<filename>* is the name of the starting image (Figure 1g, main text). All the subsequent analyses, including orientation (see below), will be performed only on such selected pixels.

At this point, the algorithm forks (Figure 1h, main text): the first branch—always performed—is dedicated to cell morphometry tout-court, while the second branch—optional—is devoted to the orientation analysis and it will be described in the next section.

MORPHEUS cell morphometry consists in the evaluation of 12 different shape descriptors for each detected cell. In particular, all the suitable shape descriptors natively provided by Fiji/ImageJ through the *Set Measurements* function have been included in MORPHEUS pipeline, namely:

- Area (A)
- Perimeter (P)
- Best fitting ellipse major axis (B)
- Best fitting ellipse minor axis (b)
- Best fitting ellipse aspect ratio ($AR = \frac{B}{b}$)
- Best fitting ellipse angle (θ)
- Circularity ($C = \frac{4\pi A}{P^2}$)
- Roundness ($R = \frac{4A}{\pi B^2}$)
- Solidity ($S = \frac{A}{convex\ hull}$)

- Feret’s diameter (or maximum caliper) (F)
- Feret’s diameter angle (φ)
- Minimum caliper diameter (or minimum Feret’s diameter) (f).

Although with a certain degree of redundancy, this set covers the most known and used shape descriptors in the field of morphometry. The idea here is to use a descriptor spectrum as wide as possible in order to capture as many shape features as possible, without the need of choosing any particular descriptor ‘*a priori*’. Then, during the follow-up analysis, it will be possible to resort to some technique of dimensionality reduction to identify the most informative minimum set of variables required by the problem (see Figure 1j in the main text and the following *Experimental Validation – Morphometry* section).

For each sample image, MORPHEUS saves a separate report file (in Microsoft Excel format) in the *Cell* output subfolder. This file is named *Descriptors_#.xls* and contains the values of the 12 descriptors evaluated for each cell identified within the $\#^{\text{th}}$ sample image ($\# = 1, 2, \dots, m$). The final MORPHEUS output is the assembly of such partial descriptor matrices, that is to say a new $n \times 12$ matrix— n being the total number of detected cells—where each row is a 12-dimensional vector containing the ordered list of the descriptors (or features) referred to a single cell (Figure 1i, main text). MORPHEUS saves this table as *MasterMatrix_M.xls* in the output folder after having removed from descriptor spectra all those non-informative variables evaluated by *Set Measurements* function by default—namely the coordinates locating the Feret’s diameter (*FeretX* and *FeretY*), the upper left corner of the *bounding rectangle* (*BX* and *BY*), and the center of the best fitting ellipse (*X* and *Y*).

If *Nucleus analysis* option was checked in the starting window, MORPHEUS also performs a separated analysis of those images featuring in their file name the *nucleus identifier* specified by the user (e.g. *dapi*). This parallel morphometric analysis consists of the same operations described above for entire cells, the only difference being an additional watershed step at the end of the segmentation procedure. The results of such a nucleus segmentation and identification are saved

in a separated subfolder—named *Nucleus*—within the output directory. In addition, MORPHEUS returns a second descriptor matrix—named *MasterMatrix_N.xls*—featuring a row for each detected nucleus. Here the first branch of the main algorithm terminates and, if enabled, MORPHEUS proceeds with the second one, namely the orientation analysis (see next section).

No further manipulation of these descriptor matrices is implemented in MORPHEUS, since, in general, it heavily depends on the particular biological question from which the experiment has originated and, for the same reason, it is often better to proceed using different analysis environments. However, even if the end user can decide to freely manage the n -by-12 matrix as he/she prefers, in the spirit of making MORPHEUS a complete and standardized pipeline for cell morphometry, we also propose a general framework for the post-processing of *MasterMatrix_M/N*. Specifically, depending on the particular biological issue, three main different procedures can be followed covering most of the situations.

(i) If the aim is testing for the presence of significant differences among $k \geq 3$ conditions, a (one-way) Multivariate ANalysis Of VAriance (MANOVA) can be conducted, that reduces to an Hotelling's T -squared test for two independent samples when $k = 2$. In this case, descriptor spectra are regarded as (continuous) dependent variables in the context of a General Linear Model (GLM) framework, where the independent categorical variables define the experimental groups. Discriminant Function Analysis (DFA) and Canonical-Correlation Analysis (CCA) are two suitable alternative approaches to this kind of problem. (ii) If the experimenter is interested in using descriptor spectra as independent variables (regressors) to infer a deterministic relationship between them and some cellular function or biological process of interest—according to the aforementioned shape-function paradigm—a Multiple Linear Regression (MLR) should be applied. In this particular GLM, both independent and dependent variables can be multidimensional, thus meaning that the biological processes representing the independent variables can be more than one at the same time (e.g. proliferation, differentiation, migration), but all of them need to be captured by some

continuous measure. Moreover, in order to successfully accomplish the regression, correlation among the shape descriptors (multicollinearity) needs to be removed by means of a preparatory Principal Component Analysis (PCA). In this regard, Partial Least Squares (PLS) regression is an effective—and probably more powerful—alternative to MLR for modelling the existing relationship between cell shape and one (or more) continuous biological process. (iii) Finally, descriptor spectra can be used as independent variables for classification purposes. Linear Discriminant Analysis (LDA)—rather than k -Nearest Neighbors (k -NN) or Support Vector Machines (SVMs)—can be used to accomplish such a task, somehow complementary to task (i). In this case, a training set of pictures/descriptors is used for a supervised learning step to generate the model that will allow the classification of new samples of unknown origin (healthy vs disease, treated vs control, etc.). Even in the case of LDA, a PCA transformation is recommended prior the regression step to eliminate multicollinearity among regressors.

In the *Experimental Validation – Morphometry* section, a detailed illustration of the first task (two-group statistical comparison) is presented using original fluorescence images.

1.4. Orientation Analysis

In addition to the assessment of the morphometric descriptors, MORPHEUS can also implement an automated directional analysis making use of OrientationJ (Daniel Sage, Biomedical Image Group - BIG, EPFL, Lausanne, Switzerland), a well-established plugin for Fiji/ImageJ [4]–[6]. OrientationJ provides a quantitative characterization of the orientation and the isotropy properties of a region of interest within an image. Specifically, starting from the partial derivatives of pixel intensity values, it computes a 2×2 matrix—called structure tensor—containing all the relevant directional information. The analysis is performed locally by means of a sliding Gaussian window with arbitrary radius σ , whose value should be chosen as close as possible to the structure of interest. Three main features can be extracted from the structure tensor: orientation (the angle of the dominant

orientation of the region), coherency (the local anisotropy, or orientation magnitude) and gradient energy (the square modulus of the local gradient). For more details on OrientationJ computational aspects, see the references above and the official documentation on the website:

<http://bigwww.epfl.ch/demo/orientation/>.

In order to make MORPHEUS able to perform directional analysis, OrientationJ plugin (version 2.0.2 or above) needs to be installed in *plugins* subfolder before opening Fiji and launching MORPHEUS. If the corresponding checkbox in the starting window of MORPHEUS has been selected, after the assessment of morphometry, the algorithm checks for the presence of OrientationJ plugin and, if present, *OrientationJ Distribution* function is repeatedly called to perform an automated orientation analysis of the whole dataset over two distinct scales, namely cytoskeleton and whole-cell level (Figure 1k, main text). For cytoskeleton analysis, original images are masked using the outlines of the isolated cells detected as a result of the primary learning step (i.e. *Img_#b - Detection - <filename>.tif*; Figure 1g in main text). By this way, background, debris and cell clusters are discarded (set to 0) and the only pixels retained are those corresponding to single cells. Such images are used as input for *OrientationJ Distribution* function and the radius of the local window employed for structure tensor computation is set to its minimum value ($\sigma = 1$ pixel) to make it suitable for fine structure analysis. On the contrary, for the whole-cell level, intracellular structure information is discarded and the binarized images of isolated cells are used as input for *OrientationJ Distribution*. In this case, in order to allow structure tensor capturing features at the cellular scale, the radius of the Gaussian window is set to a value proportional to the square root of the ‘typical cell area’ as estimated in the learning step (see main text).

For both the levels of analysis—cytoskeleton and whole-cell—MORPHEUS returns different qualitative and quantitative outputs. In particular, two colormaps are saved in *Cell* output subfolder as *Img_#c - CellOrient - <filename>.tif* and *Img_#d - CytoOrient - <filename>.tif*, respectively. These images carry the

directional information according to a Hue-Saturation-Brightness (HSB) color model, where hue encodes for the local dominant orientation (the direction along which elements are orientated), saturation encodes for coherency (an estimation of how much they are elongated in that direction) and brightness is based on the gray levels of the input image (Figure 1l, main text). In addition, a set of coherency-weighted orientation histograms are collected and then assembled into a single *.tif* heatmap called *MasterMatrix_O*. Such a matrix is displayed when MORPHEUS terminates (Figure 1m, main text) and then saved in the form of both image and table (Excel format) in the output directory. *MasterMatrix_O* is a $180 \times 2m$ matrix (m being the total number of sample images) in which each row corresponds to a particular orientation in the plane (from $+89^\circ$ to -90° , in steps of 1°), while each column pertains to a different sample image, the first m columns being assigned to whole-cell histograms and the following ones to cytoskeleton (with an empty row separating the two submatrices). Orientation data are therefore presented in binned form (by a binning width of 1°) and finally renormalized to n (the total number of isolated cells detected) to make them suitable for subsequent statistical hypothesis testing (see below). In this regard, it is worth noting that the selection of the sole isolated cells operated by MORPHEUS in its preliminary step represents a way to control the number of independent angular observations for properly calculating *p*-values (more details about this point can be found in *Experimental Validation – Orientation Analysis* section).

Notably, to restrict the analysis to the pixels actually belonging to the detected cells and avoid any interference from the surrounding regions, a ‘tweak’ of OrientationJ algorithm was required. In particular, MORPHEUS forces OrientationJ to use the outlines of the detected cells (Figure 1g, main text) as a selection mask to overcome the poor selection capabilities of the original plugin. OrientationJ offers indeed the possibility to manually select individual objects for subsequent orientation analysis (see *OrientationJ Measure* function), but this feature poorly matches the highly automated approach of MORPHEUS. Alternatively, OrientationJ provides some automated masking options in

OrientationJ Distribution function that are based on coherency or energy local values. However, such a selection process is subjected to the constriction of using the same spatial resolution ($\approx \sigma^{-1}$) employed for the detection of the dominant orientation, which could be fine in the case of cytoskeleton ($\sigma = 1$ pixel, implying the maximum spatial resolution), but is largely inadequate for whole-cell analysis, especially with sparse cell culture. Importantly, the tweaking of OrientationJ capabilities is operated by MORPHEUS at runtime on the output of the original plugin: for this reason, there is no need to download any ‘modified’ version of OrientationJ.

Here MORPHEUS terminates, since, as in the case of morphometry, the follow-up analysis depends on the specific information the researcher is interested in extracting from the data. A first qualitative information can be retrieved simply by the observation of *MasterMatrix_O*, because any possible overrepresented (or *enriched*) direction would appear as a horizontal warmer band in the heatmap. Notice also that the two submatrices making up *MasterMatrix_O* usually have a quite different appearance, in that the one on the left (cellular scale) generally exhibits ‘grainier’ and more discretized patterns if compared to the right one (cytoskeleton scale), which shows instead diffused distributions of orientation values. This clearly reflects the different number and size of the coherent objects considered at the two different scales: a relatively small number of whole-cells in the first case and a huge number of tiny filaments in the latter.

However, to move from the qualitative to a quantitative level it is typically necessary to resort to the circular statistics. The basic analysis may consist in some statistical procedure to test the resulting orientation distributions against the null hypothesis of uniformity, making it possible to determine if cells are randomly oriented without any preferential direction, or, on the contrary, if there exists some significant deviation from the uniform distribution (Figure 1n, main text). This may be very useful for assessing the orientating power of particular substrates rather than the attractive effect exerted by the gradient of some chemoattractant. To this purpose, a number of statistical descriptors and hypothesis tests

specifically designed for circular data—namely the Rayleigh, Rao’s spacing and V-tests—exist in literature [7]–[10]. An example of such an analysis is given in *Experimental Validation – Orientation Analysis* section, where the steering power of a machined titanium substrate is investigated using original fluorescence images.

Finally, it is to be noticed that, even if the two pieces of information conveyed by the cellular and cytoskeletal submatrices are typically highly correlated because of biological reasons, from a computational point of view they are completely independent and the user can choose which of the two levels is the most informative within that particular experimental context (e.g. the fine scale could be preferred for investigating cytoskeleton integrity, while whole-cell analysis is suitable for studying the polarization of the cells or in case of cytosolic dyes; see *Angular Resolution and Scale Dependency* section for a demonstrative test of such an independence).

1.5. MORPHEUS Log

With the aim of being as compliant as possible with the reproducibility standards in bioinformatics (see e.g. [11]), MORPHEUS prints an onscreen log at runtime and, when the algorithm terminates, a copy of this log is saved as a `.txt` file in the output directory. Both the onscreen log and the log file contain useful information for reproducing that particular analysis at a later time, with a possible different informatics architecture. This information encompasses the particular values chosen by the user for the two parameters ϵ and T , the architecture of the system, the OS type and the version of Java, Fiji/ImageJ and MORPHEUS itself. MORPHEUS also checks for the presence of those third-party plugins that are essential for its proper functioning (i.e. *Auto Threshold*, *OrientationJ*) and reports their version in the log. In addition, MORPHEUS log returns many messages about the picture-file currently being processed and, if *Verbose log* option was checked in the starting dialog window, it also shows the values of area and circularity computed during the unsupervised learning and then used to filter out those

objects that are not isolated cells. The number of detected (segmented) objects and the size n of the subset of identified cells are also printed. Finally, if everything worked properly, the message “All output files have been correctly saved to <output>” is printed.

2. Statistical Validation

In *Sample Preparation* section in the main text, some recommendations are made to make MORPHEUS able to effectively distinguish isolated cells from cell clusters after the segmentation process. In particular, moderate and homogeneous plating densities are supposed to provide in each field (or sample image) a majority of isolated cells compared to cell clusters (point 4) and more than 10-15 cells per field are recommended as sample size (point 6). The first requirement is necessary for median area being a good sample statistic in representation of the subpopulation of isolated cells, reducing the influence of outliers and cell clusters. The second requirement is needed instead to have a more stable sample median and to produce a low-skewness sampling distribution, for which mean and standard deviation can be considered good measures of central tendency and dispersion respectively. Here we intend to provide some theoretical justification regarding the particular recommended sample size.

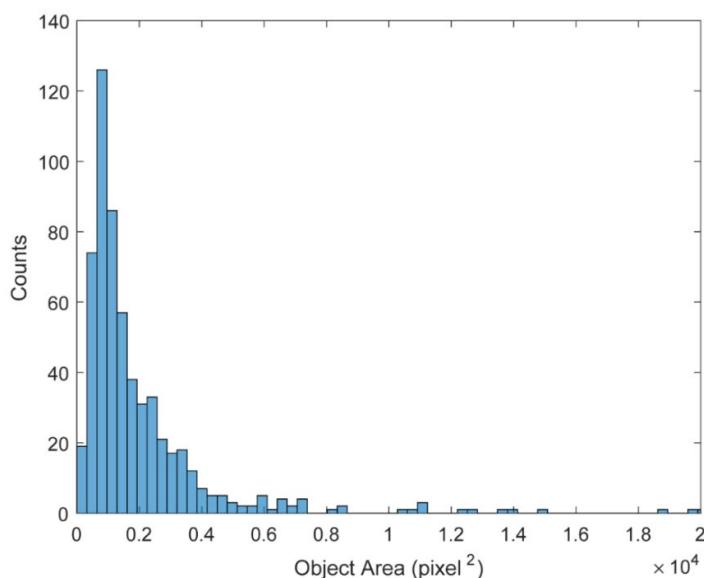


Figure SM2. Histogram of areas. Areas of all the objects segmented by MORPHEUS during the preliminary analysis (learning step) of PCC dataset, one of the three experimental groups used for the following validation of morphometric analysis.

Under the hypothesis that isolated cells are indeed more abundant than cell clusters, the distribution of the areas of all the segmented objects will tend to be unimodal, with a main low-valued body—single cell subpopulation—and a right

heavy tail, given by subpopulations of clusters of two or more cells. This intuition was confirmed by empirical data: Figure SM2 shows a representative histogram of the areas as returned by the preprocessing of a real experimental dataset. This particular set of images was denoted as PCC and it has been also used for morphometry validation (see *Experimental Validation* section for a detailed description of such a cellular model and dataset).

To fit such histograms, a number of popular continuous probability distributions may be considered, including exponential, gamma, Weibull, log-normal, log-logistic and Burr distribution. All these six models were tested by fitting the data through maximum likelihood estimation, in each one of the three conditions used for the validation of morphometric analysis (EC+, EC-, and PCC, see *Experimental Validation* section below). Best-fit parameters were then used for judging the goodness of fit of the six theoretical models compared to the empirical data. Test results, expressed in terms of p -values, are reported in Table SM1, where upper values refer to one-sample Kolmogorov–Smirnov test, while lower ones to one-sample Cramer–von Mises test.

Table SM1

Distribution	EC+	EC-	PCC
<i>Exponential</i>	0.0062	0.1001	$1.07 \cdot 10^{-7}$
	0.0047	0.0976	$< 10^{-4}$
<i>Gamma</i>	0.1997	0.1744	$3.91 \cdot 10^{-9}$
	0.1110	0.1239	$< 10^{-4}$
<i>Weibull</i>	0.2627	0.1814	$1.00 \cdot 10^{-7}$
	0.2207	0.1262	$< 10^{-4}$
<i>Log-Normal</i>	0.0767	$4.36 \cdot 10^{-4}$	0.0211
	0.0740	$8.33 \cdot 10^{-4}$	0.0486
<i>Log-Logistic</i>	0.0454	0.0237	0.3893
	0.1770	0.0234	0.1911
<i>Burr</i>	0.2357	0.1602	0.4123
	0.3348	0.1218	0.5910

While ‘the best’ theoretical model underlying these highly right-skewed populations probably depends on the particular experimental configuration (cell type, plating conditions, objective magnification, etc.), in general we observed that right tails were typically heavier than log-normal. Ultimately, log-logistic distribution proved to be a good compromise between fitting performance and parametric simplicity (Figure SM3). Only two parameters (μ and σ) are indeed needed for defining a log-logistic probability density function f (while Burr distribution relies on three) and both of them showed a reduced range of variability when applied to our experimental data (see Table SM2). Specifically, the log-logistic PDF can be parametrized as follow:

$$f(x | \mu, \sigma) = \frac{1}{\sigma e^\mu} \frac{\left(\frac{x}{e^\mu}\right)^{\frac{1}{\sigma}-1}}{\left[1 + \left(\frac{x}{e^\mu}\right)^{\frac{1}{\sigma}}\right]^2} = \frac{1}{\sigma x} \frac{e^z}{(1 + e^z)^2} ; \quad x \geq 0, \text{ where } z = \frac{\ln(x) - \mu}{\sigma} .$$

Table SM2

Parameter	EC+	EC-	PCC
<i>Log-median - μ</i>	7.90	8.06	7.18
<i>Shape parameter - σ</i>	0.78	0.68	0.47

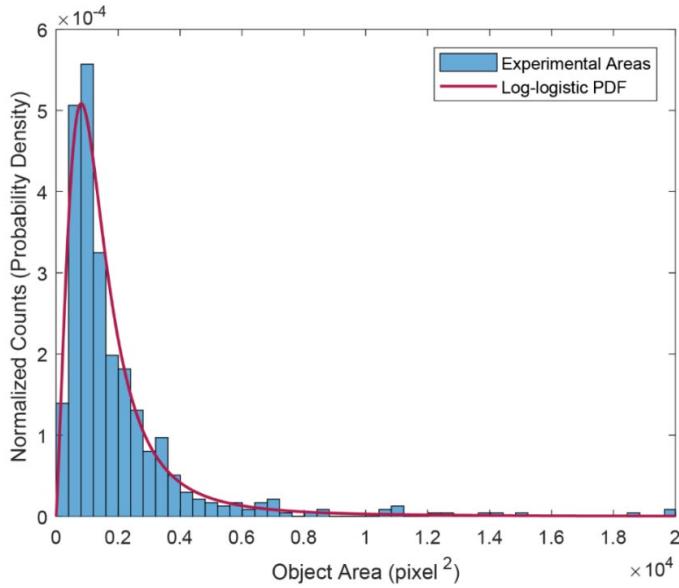


Figure SM3. Experimental data and theoretical model. PCC areas can be satisfactorily fitted by a Log-logistic PDF: $\mu = 7.18$, $\sigma = 0.47$. p -value = 0.39, one-sample Kolmogorov–Smirnov test.

Importantly, since MORPHEUS is based on a completely non-parametric approach, we are not interested here in finding the most reliable theoretical model underling the distribution of areas. Rather we want to use a plausible model—as the log-logistic one—to show how the skewness of the sampling distribution of sample median decreases when increasing sample size. Furthermore, we want to find a reasonable cutoff for such a distribution to be considered sufficiently symmetric. For this, we performed a Monte Carlo simulation to evaluate the skewness of the sampling distribution of sample median as a function of sample size p , using a log-logistic underlying distribution for different values of σ . Specifically, we generated $m = 10^6$ log-logistically distributed samples of size p , computed the median for each one of them and finally evaluated the skewness of the resulting sampling distribution as a function of p (from 1 to 50). This was done for four realistic dispersion values σ (from 0.5 to 0.8, in steps of 0.1) and for a fixed location value $\mu = 8$. Results presented in Figure SM4 clearly show that—depending on the dispersion of the underlying distribution—a sample size ranging between 10 and 25 can usually ensure an almost symmetric ($\text{skewness} < 1$) sampling distribution of the sample median. Thus, mean and standard deviation can be conveniently used to measure the

central tendency and dispersion, respectively, of such a sampling distribution. This is actually the method MORPHEUS uses to learn the *characteristic area of the single cell* and its uncertainty. Upon these estimates, the *single cell size range* is consequently defined (see *Algorithm Flow – Unsupervised Learning Step* section in the main text).

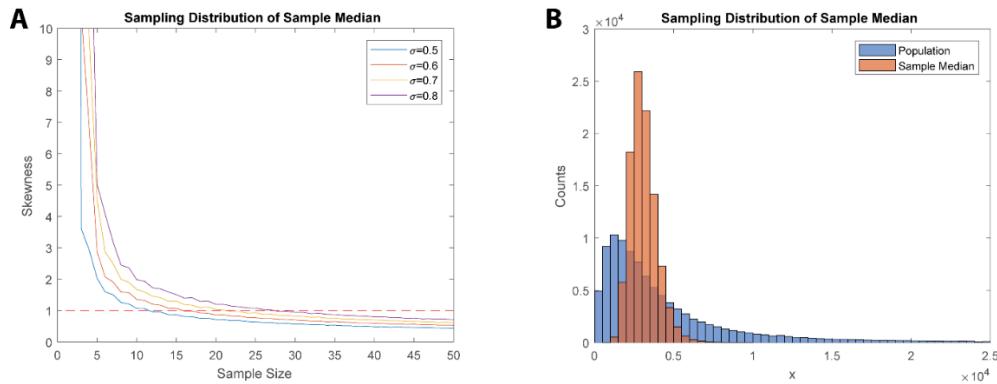


Figure SM4. Analysis of the sampling distribution of the sample median through Monte Carlo simulation. (A) Skewness of the sampling distribution of the sample median as a function of sample size p for four different dispersion values σ of the underlying log-logistic distribution (10^6 random samples for each point and $\mu = 8$ in all cases). (B) A representative comparison between a log-logistic population of parameters $\mu = 8$ and $\sigma = 0.6$ (blue bars) and the related sampling distribution of the sample median (red bars) with sample size $p = 20$ (10^5 random samples were generated for both the histograms).

It is worth noting that these results were not obvious *a priori*, especially if considering the impressive instability of the sample mean and the extremely slow rate by which it approaches to normality when the underlying distribution is log-logistic—in particular when σ is close to 1—despite the central limit theorem (CLT). This is due to the extreme heaviness of the right tail of the log-logistic distribution, for which the mean is not even defined when $\sigma \geq 1$, such as in the case of other notable distributions (e.g. the Pareto distribution). In this regard, Figure SM5 shows the results of another Monte Carlo simulation featuring the same parameters of the previous one, but in which the mean was sampled in place of the median, for sample size p ranging from 1 to 100.

In summary, we proved that—in the case of a log-logistic-like underlying distribution—the median is a suitable statistic to be sampled because of the stability and fast decay of the standardized moments of its sampling distribution, for increasing sample size p . In particular, for $p > 20$, sampling distribution of the

sample median is symmetric enough to be conveniently described by its mean and standard deviation, thus providing a reliable estimate of population median and its uncertainty.

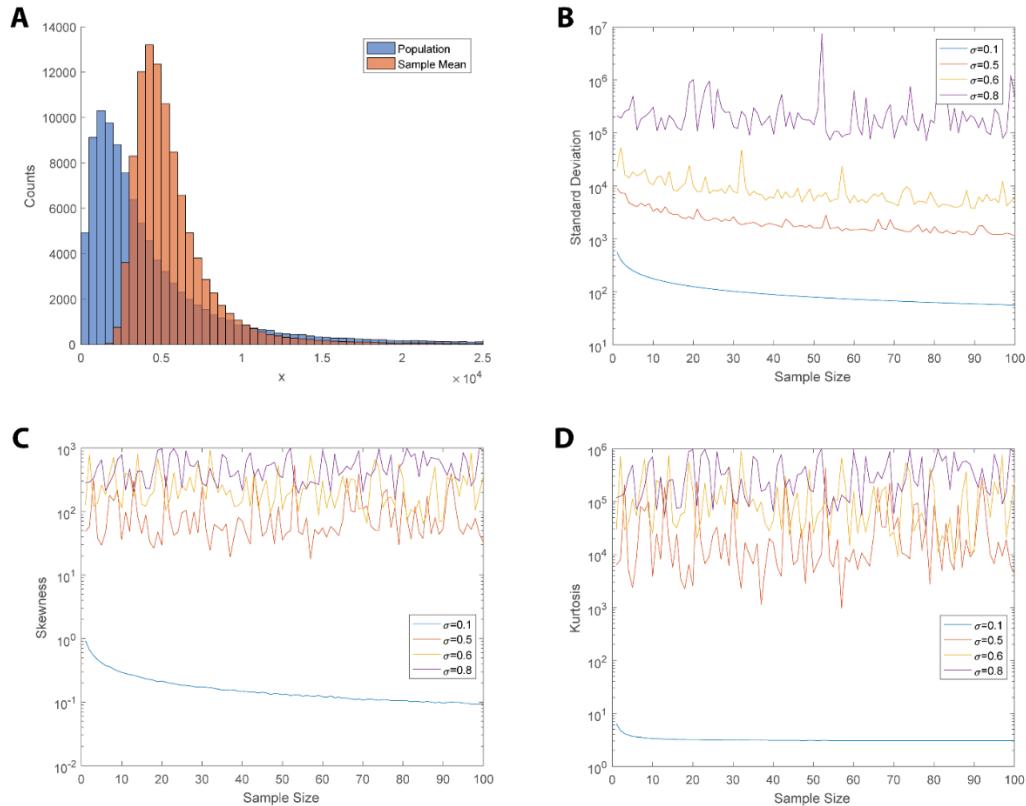


Figure S5. Analysis of the sampling distribution of the sample mean through Monte Carlo simulation. (A) Comparison between a log-logistic population of parameters $\mu = 8$ and $\sigma = 0.6$ (blue bars) and the related sampling distribution of the sample mean (red bars) with sample size $p = 20$ (10^5 random samples were generated for both the histograms). Notice the high skewness and the heavy right tail still carried by the sampling distribution of the sample mean compared to those of median (Figure SM4B). **(B, C, D)** In the order: Standard deviation, skewness and kurtosis of the sampling distribution of the sample mean as a function of sample size p for four different dispersion values σ of the underlying log-logistic distribution (10^6 random samples for each point and $\mu = 8$ in all cases). Only for σ values much smaller than the experimental ones, CLT ‘normalizing effect’ (skewness $\rightarrow 0$ and kurtosis $\rightarrow 3$ as $p \rightarrow \infty$) could be easily visualized. Otherwise, central and standardized moments showed great instability and very slow rates of convergence.

3. Experimental Validation

3.1. Cell Cultures and Image Acquisition

For the experimental validation purpose, Human Microvascular Endothelial Cells (EC), PC3 cells (PCC) and Human Dermal Fibroblast (HDF) were used [12]–[14]. These cells were maintained respectively in the following growth media: EndoGRO-MV (Millipore); RPMI 1640 (Gibco) supplemented with 10% Fetal Bovine Serum (FBS) (Gibco) and DMEM (Gibco) supplemented with 10% FBS. All media were supplemented 100 U/ml penicillin, 100 µg/ml streptomycin and L-Glutamine 10 mM (Sigma Aldrich). Cells were passaged at subconfluence to prevent contact inhibition and were kept under a humidified atmosphere of 5% CO₂ in air, at 37 °C.

Cells were seeded at a concentration of 2000 cells/well in a 24-well plate on different substrates. In particular, polystyrene plate, polystyrene plate with 1% gelatin coating or machined titanium disks (characterized by concentric grooves due to the production process) [15]–[18] were used. After 24 hours, cells were fixed in 4% paraformaldehyde in Phosphate Buffer Saline (PBS) and stained with Rhodamine-Phalloidin and DAPI (Life Technologies, Milan, Italy) to highlight actin network and nuclei respectively [19]–[21].

Images were acquired with a Nikon Eclipse Ti-E microscope using different objectives: Nikon Plan 10x/0.10 (EC and PCC); Nikon Plan Fluor 20x/0.50 (HDF); Nikon Instruments.

3.2. Morphometry

Among the three tasks outlined at the end of the previous *Morphometry Analysis* section, the first one is probably also the simplest one, especially when $k = 2$. Here, to provide a first and straightforward experimental validation of MORPHEUS, two examples of such an analysis are presented.

As a first validation trial, we started comparing two experimental groups consisting in cell cultures exhibiting markedly different morphometric features, in

order to focus on the meaning of MORPHEUS output and present an effective way to retrieve information from it. In particular, we tested endothelial cells plated with a complex extracellular matrix coating (hereafter EC+) against a prostate cancer PC-3 cell line (hereafter PCC). Figures SM6A and SM6B show two sample images representative of EC+ and PCC groups respectively. Specifically, EC+ group consisted of $m = 14$ sample images (from which MORPHEUS was able to detect $n = 133$ isolated cells out of 232 segmented objects), while PCC group was made of $m = 30$ sample images (from which MORPHEUS detected $n = 431$ isolated cells out of 591 segmented objects). Starting parameters have been set to $T = 5$ and $\varepsilon = 200$ pixel² for all the analyses presented in this section. However, the influence of the tolerance parameter on the number of cells identified by MORPHEUS is detailed in Table SM3 of the following *Tolerance Effect* section.

For each one of the two sets of images, EC+ and PCC, MORPHEUS returned *MasterMatrix_M.xls* file that is a $n \times 12$ matrix containing (for each detected cell) the evaluation of the 12 morphometric descriptors already presented. In order to provide a visual representation of the data, a preliminary PCA was conducted to reduce the dimensionality of the problem. Notice that the different shape descriptors evaluated by MORPHEUS have, in general, very different domains and variability ranges: for instance, circularity, roundness and solidity values are bounded between 0 and 1, while area and perimeter can vary from 0 to infinite, with the former typically exhibiting numerical values (in pixel) much larger than the latter. Importantly, since PCA is a scale-sensitive method, raw data need to be standardized (z-scored) before the covariance matrix is diagonalized or, that is the same, PCA needs to be conducted by diagonalizing the correlation matrix. Specifically, the scree plot in Figure SM6C shows how the first 3 principal components (PCs) were indeed able to explain about the 90% of the total variance of the EC+/PCC dataset (thus confirming the high degree of correlation among the 12 original morphometric descriptors). The marked separation between the two groups can be clearly observed in Figure SM6D, showing the scatter plots of the transformed data (also known as *score plots*) in the 3-dimensional space spanned

by the first 3 PCs. Moreover, it is important to notice how only the first PC (PC1) actually described the *between-group* variance, while PC2 and PC3 just seemed to carry information about two *within* components of the total variance. However, in this type of analysis, PCA is just a helpful tool for the visual inspection of an otherwise non-representable 12-dimension dataset, but it is not a strictly necessary step, since MORPHEUS output is indeed directly suitable for statistical testing. Specifically, a Hotelling's *T*-squared test for two independent samples (not assuming equal covariance matrices) was performed on the original dataset, leading to an extremely low p -value = $1.73 \cdot 10^{-256}$ and confirming that a significant difference existed between the morphology of the cells belonging to the two groups. In order to investigate which features in particular accounted for such a difference, a series of follow up procedures after the Hotelling's omnibus test were conducted. First, the average standardized descriptor spectra (i.e. the two within-group means of the z-scored original data) were plotted to look for the descriptors at which the two spectra differed most (Figure SM6E, upper). To this purpose, an effective way to plot such a distance could be the computation of Cohen's d effect size, a measure that takes into account both mean values and their uncertainty in terms of (pooled) standard deviation (Figure SM6E, lower). Afterwards, 12 univariate *t*-tests were performed between the marginal distributions of each descriptor and the resulting p -values were used as ranking index to spot the most influencing variables (Figure SM6E, lower). In this case, both the procedures agreed to identify circularity, solidity, Feret's diameter, roundness and perimeter as the most important variables to which the difference between the groups could be ascribed. On the contrary, best fitting ellipse minor axis and angles were the least significant descriptors in terms of difference between the two subsets. These findings are reasonable and were also largely predictable since one of the main differences that could be easily observed—simply by eyeballing the sample images—was that PCC cells had a much more circular shape if compared to EC+. On the other hand, since cells had not been subjected to any orientation stimulus,

angles were not expected to be a distinctive feature, being all directions equally probable.

Importantly, this approach is straightforward but it also has many drawbacks, especially when dealing with smaller effect sizes. In fact, it is not uncommon to happen that, even in the presence of a significant multivariate Hotelling's T -squared test, none of the marginal univariate t -tests show a significant p -value. This is because univariate tests do not take into account the correlation among the variables—as the multivariate tests do—and when no significance is detected among the marginal distributions this simply means that no one of the dependent variables alone is able to account for the difference between the two groups. Rather, only the combination of two or more of them can explain the effect. Another disadvantage arising from ignoring correlation during the follow up investigation, is the possible high degree of redundancy among the descriptors retained because of their small marginal p -value (or their large marginal effect size). This is quite evident also in this particular case, being circularity, roundness, and solidity often highly correlated. For all of these reasons, an alternative and more rigorous approach consists in the so-called *loading analysis* after PCA. Loadings can be viewed as the weighting coefficients that must be used to mix the original variables in order to obtain the principal components. They can be represented as curves in the shape descriptor domain (Figure SM6F), rather than together with the scores in a so-called *biplot*, where loadings are seen as the projections of the 12-dimensional original frame of reference in the 2- or 3-dimensional space of the principal components (Figure SM6G). Notably, in this case the PC1 loading profile was almost identical to the average EC+ mean spectrum, thus confirming that PC1 was the only component actually needed to model the *between-group* variance, while PC2 (mainly determined by fitting ellipse minor axis) and PC3 (mainly determined by angles) only accounted for minor *within* variance components. Loading axes in the *biplot* were in agreement with these observations. In addition, the *biplot* made it evident how Feret's diameter and fitting ellipse major axis were positively correlated descriptors

(loading axis were very near; $\rho = 0.95$) and both of them were negatively correlated with circularity and solidity (whose axes lied on the same direction but in the opposite quarter). In general, correlation matrix showed absolute values greater than 0.8 between any of these four descriptors, that is to say they carried approximately the same information about the morphometric difference between EC+ and PCC. Finally, this finding suggested that it was presumably unnecessary to retain them all for a possible subsequent analysis. Specifically, since Feret's diameter showed the greatest influence over the PC1, it could be conveniently regarded as the best candidate to characterize the effect under examination, even by a univariate statistic.

In general, in order to accomplish the task of identifying the most influencing descriptors, without overlooking the complexity of the multivariate statistical analysis, a consensus-based approach of this kind is recommended as best practice.

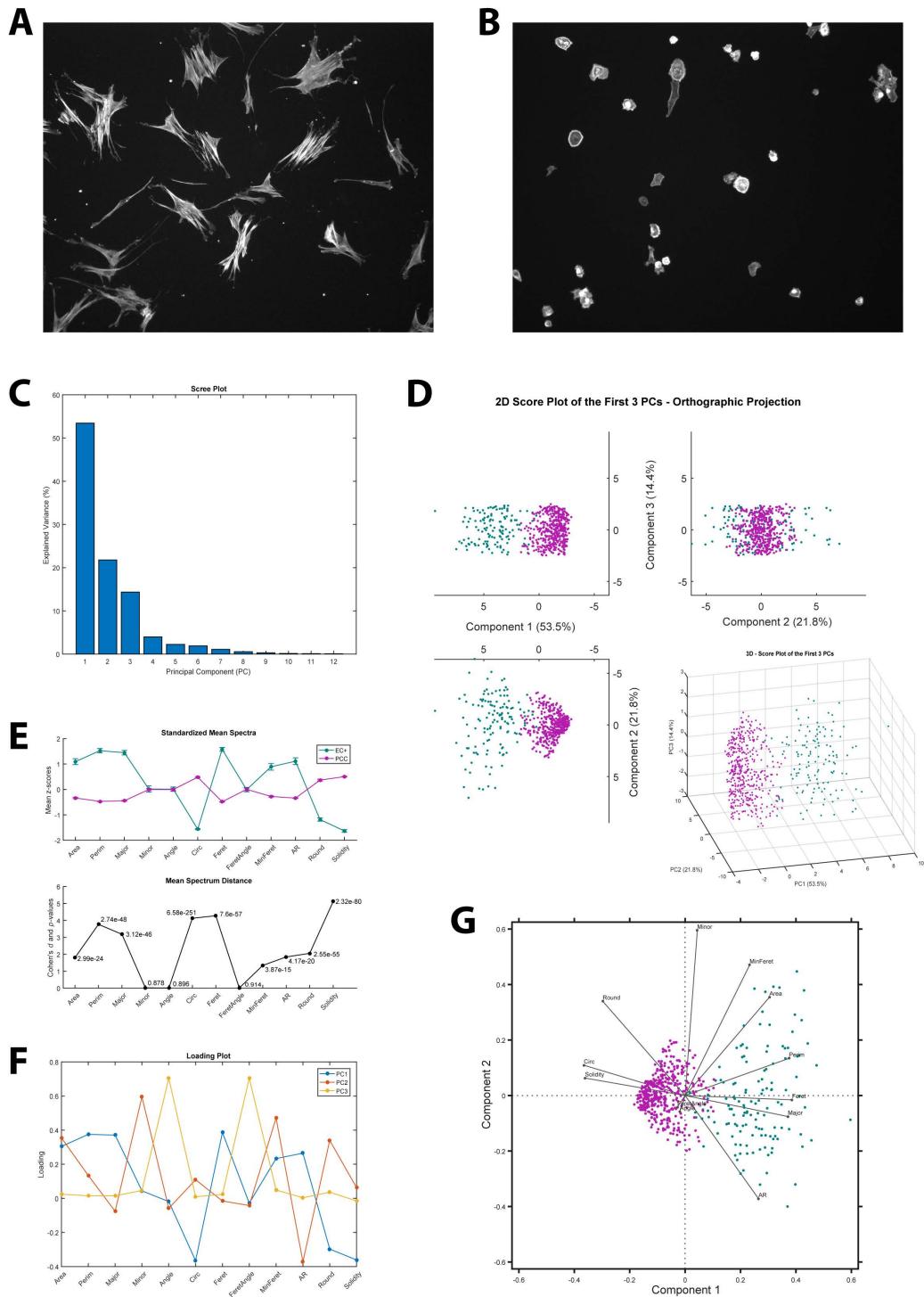


Figure SM6. Follow-up analysis of morphometry results from EC+ vs PCC comparison. **(A)** A representative image (out of 14 acquired fields) of endothelial cells plated after dish coating and fluorescently stained for cytoskeleton through phalloidin (EC+ experimental group). **(B)** A representative image (out of 30 acquired fields) of prostate cancer tumor cell line PC-3 fluorescently stained in the same manner (PCC experimental group). **(C)** PCA scree plot representing the percentage of variance explained by each principal component. The first three PCs alone account for the 89.6% of the total variance of the dataset. **(D)** PCA score plot in the space of the first three PCs: orthographic projection and 3D representation in the lower right corner. Each point (score) represents a single cell detected by MORPHEUS after its learning step (here and hereafter: EC+ in teal and PCC in magenta). The two groups appear perfectly separated as a consequence of their extremely divergent morphometric features. In particular, it is evident that the between-group variance is entirely lying

along PC1, while PC2 and PC3 just captured within-group components. **(E)** Analysis of descriptor marginal distributions. Upper plot: comparison of the two average (z-scored) descriptor spectra: for each descriptor, the distance between the curves is a measure of its contribution to the overall separation between the groups. Lower plot: Cohen's d computed for each descriptor is a suitable metric to capture such a distance. Next to each point is the p -value resulting from a t -test conducted over the marginal distribution of each descriptor. As expected, the lowest p -values are in the correspondence of the highest Cohen's d values (even if the ranking is not the same). **(F)** Comparison of the loading profiles of the first three PCs (i.e. the contribution of each descriptor to that particular PC). PC1 loading curve (in blue) is almost identical to the profile of the average descriptor spectra, thus confirming that the first PC alone is sufficient to explain the between-group variance of the entire dataset. **(G)** Biplot (i.e. scores and loadings within the same plot) in the space of the first two PCs. The choice of the descriptor closest to PC1 (i.e. Feret's diameter) allows modeling the between-group variance component by means of a univariate statistic. Other descriptors corresponding to parallel or antiparallel loading axes can be discarded because of their strong correlation (positive or negative respectively) with the former.

As a second validation test, we focused on MORPHEUS sensitivity, trying to capture the difference between two very similar experimental conditions in a less obvious and more realistic scenario. In particular, we tested the EC+ cell cultures—already used in the previous paragraph—against the same endothelial cells plated in the absence of any coating (EC- condition: $m = 16$ sample images, $n = 118$ isolated cells out of 190 segmented objects). Figures SM7A and SM7B show two sample images representative of both the conditions to better elucidate how similar the two groups were in appearance. The analysis followed the same workflow described for EC+/PCC comparison and the results are summarized in Figure SM7. In this case, the scree plot (Figure SM7C) was a little less steep than the previous one and, because of the reduced difference between the samples, the first 3 PCs could explain just the 81% of the total variance of the dataset. In fact, looking at the scores plots in Figure SM7D, there was a larger overlap between the two populations and, even in the space of the first 3 PCs, it was neither obvious how to separate the two groups, nor to understand whether a significant difference actually existed between them. Nevertheless, a Hotelling's T -squared test (two independent samples, unequal covariance matrices) returned a far greater but still highly significant p -value = $3.47 \cdot 10^{-8}$. The average standardized descriptor spectra (Figure SM7E, upper), such as the analysis of the marginal t -test p -values and Cohen's d (Figure SM7E, lower), agreed in pointing at fitting ellipse minor axis, circularity, AR, roundness, and solidity as the most important variables responsible for the *between-sample* variance. Again, angles (but in this case also perimeter, fitting ellipse major axis and minimum Feret's

diameter) were not distinctive features of the two subsets. Interestingly, the *biplot* in Figure SM7G showed that two non-significant features such as the minimum Feret's diameter and fitting ellipse major axis could be able to represent alone PC1 and PC2 respectively. This was confirmed by the analysis of the highest peaks in the corresponding loading plot profiles (Figure SM7F). On the contrary, all the aforementioned most discriminant features lied along diagonal directions in the PC1-2 space of Figure SM7G, thus suggesting that, unlike the previous case, the difference between the two samples was not captured by one of the first two PCs alone, but rather by a linear combination involving both of them. Looking back on Figure SM7D in the light of these considerations, it seems indeed possible to identify a polarization of the two groups along the diagonal northwest-southeast direction in the PC1-PC2 subspace (linear discriminant analysis could confirm this separation). Finally, the analysis of loadings in the *biplot* (Figure SM7G) suggested also a strong correlation between roundness and AR (same direction in the opposite quarter, $\rho = -0.76$) and between solidity and circularity (close axes, $\rho = 0.76$). Therefore, two out of these four descriptors could be discarded without any significant loss of information about the *between* variance component. For instance, AR and solidity could be used for a bivariate exhaustive description of the two groups within the dataset. The fact that these two variables were almost uncorrelated ($\rho = -0.06$)—and hence both of them were needed—was reflected by their reciprocal position in the *biplot*, given that uncorrelated variables tend to preserve their orthogonality even after projection in the space of the first PCs. Specifically, AR—as well as the opposite of roundness—is higher in EC+ cells compared to EC- (Figure SM7E). At the same time, solidity—and circularity as well—is higher in EC- cells (Figure SM7E). Taken together, these two observations allow to conclude that the extracellular matrix coating was able to induce a morphological change in this endothelial model, EC+ cells having on average a more oblong shape (higher AR) and more protrusions (lower solidity) than EC- cells.

In conclusion, thanks to MORPHEUS and the follow up analysis here described, the subtle effect of a coating on cell shape could be assessed and a reduced set of variables could be identified to comprehensively describe the related morphological differences.

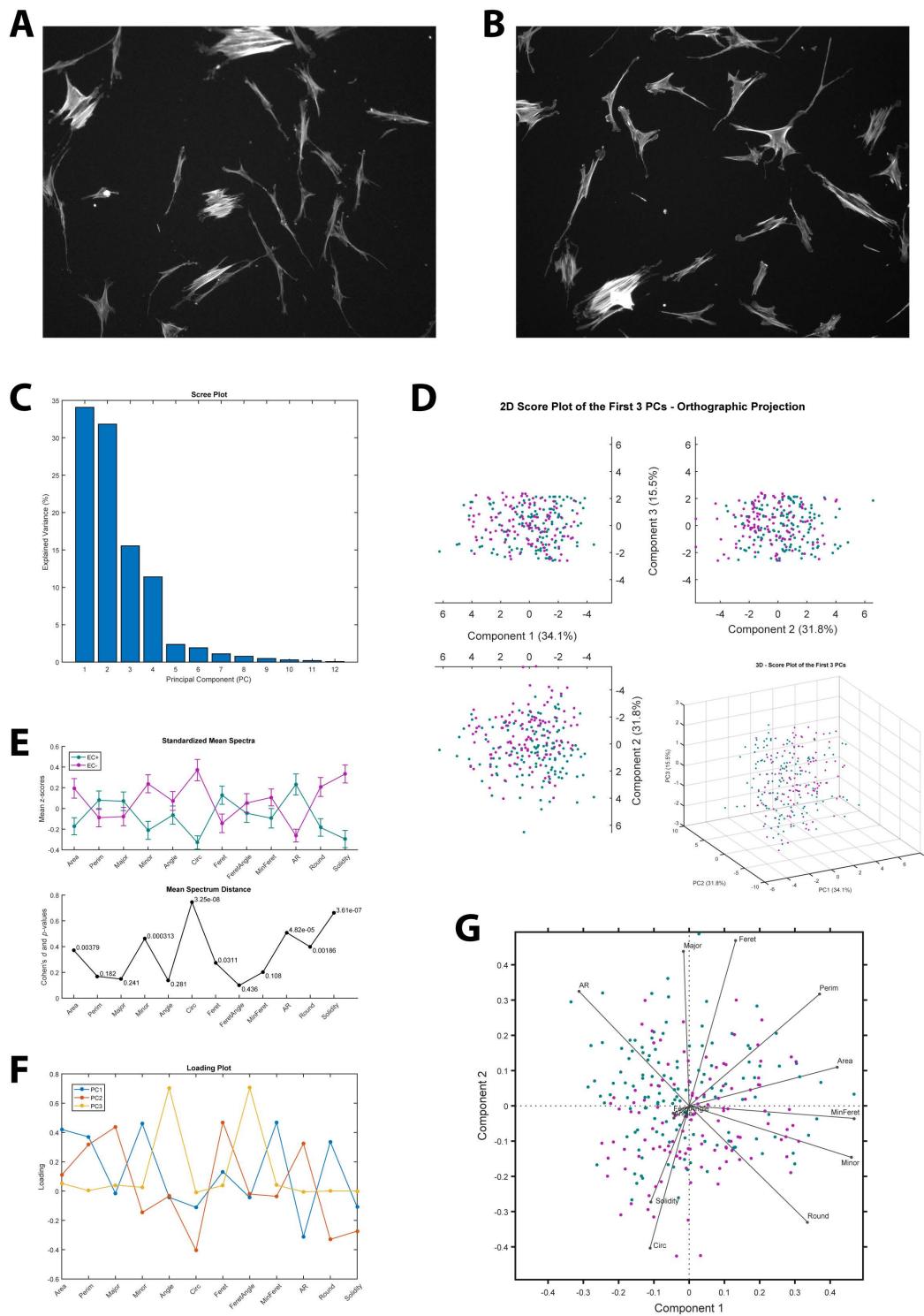


Figure SM7. Follow-up analysis of morphometry results from EC+ vs EC- comparison. **(A)** A representative image (out of 14 acquired fields) of endothelial cells plated after dish coating and fluorescently stained for cytoskeleton through phalloidin (EC+ experimental group). **(B)** A representative image (out of 16 acquired fields) of the same cell type plated in absence of dish coating and fluorescently stained in the same manner (EC- experimental group). **(C)** PCA scree plot representing the percentage of variance explained by each principal component. The first three PCs account for the 81.4% of the total variance of the dataset. Four PCs are needed to reach 92.8%. **(D)** PCA score plot in the space of the first three PCs: orthographic projection and 3D representation in the lower right corner. Each point (score) represents a single cell detected by MORPHEUS after its learning step (here and hereafter: EC+ in teal and EC- in magenta). The two groups appear to be highly

overlapping due to their subtle morphometric differences and it is not clear which PC can better model the possible between-group variance component. (E) Analysis of descriptor marginal distributions. Upper plot: comparison of the two average (z-scored) descriptor spectra: for each descriptor, the distance between the curves is a measure of its contribution to the overall separation between the groups. Lower plot: Cohen's d computed for each descriptor. Notice the general low values of this effect-size metric when compared to those referred to the EC+/PCC groups (Figure SM6E). Next to each point is the p-value resulting from a t-test conducted over the marginal distribution of each descriptor. (F) Comparison of the loading profiles of the first three PCs (i.e. the contribution of each descriptor to that particular PC). In this case, neither PC1 nor PC2 alone are able to reproduce the profile of the average descriptor spectra; on the contrary, a linear combination of the two is necessary to model the between-group variance. (G) Biplot (i.e. scores and loadings within the same plot) in the space of the first two PCs. In this frame, linear combinations of PC1 and PC2 correspond to diagonal loading axes. The choice of an orthogonal pair of such diagonal descriptors (e.g. AR and solidity) makes it possible to use a reduced bivariate statistic to describe the differences between the two groups.

Even if measuring very different effect sizes (Mahalanobis Distance $D^2 = 45.4$ for EC+ vs PCC; $D^2 = 1.1$ for EC+ vs EC-), MORPHEUS succeeded in finding statistically significant differences, with very low p -values, in both the comparisons. Since this happened even in the presence of an apparent large overlap between the two groups under examination, an excess of statistical sensitivity could be suspected. In particular, one could ask if this method is prone to attribute significant p -values even in the presence of biologically irrelevant effects, arising randomly because of the heterogeneity of cell populations. To answer this question, we split PCC dataset into two randomized subsets of 15 images each and ran MORPHEUS on both. The resulting *MasterMatrix_M* outputs were statistically tested for difference through a Hotelling's T -squared test which returned a non-significant p -value = 0.43. This further confirmed the reliability and the power of this framework of analysis.

As a final note, it should be pointed out that it is not guaranteed that PCA actually succeeds in locating the directions that best highlight the separation between the groups to be tested. However, in practice, PC1-2 are often good approximations of this, in particular when some effect exists, thus making the *between* variance component being greater than the *within* one. In general, Canonical Correlation Analysis (CCA) and Discriminant Function Analysis (DFA) are two more refined techniques that are specifically tailored for such a task.

In any case, in the presented examples, neither the optimum class separation, nor the PCA are necessary preprocessing steps, since the Hotelling's T -squared test can be conducted directly on the original variables. PCA just allowed reducing the

dimensionality of the problem by removing redundancy from the descriptors. This in turn facilitated the visualization and the understanding of the data, made it possible to assess the number of effective dimensions and the original features that were most relevant for the problem. Conversely, in the case of MANOVA (three or more groups tested at the same time) it is known that statistical power is affected by the correlation of the dependent variables [22]. In this case, when a high degree of correlation ($\rho > 0.9$) is detected between two or more variables, PCA-preprocessed data should be used for statistical analysis.

3.3. Tolerance Effect

For each one of the three experimental groups presented in the previous paragraph, the following Table SM3 shows, in the order: the mean value of M_A , the related standard error $SE_{\bar{M}}$, the ranges of A and C used by MORPHEUS to filter isolated cells, the number of identified cells, and the number of segmented objects as a function of the tolerance parameter T . As expected, both \bar{M}_A and $SE_{\bar{M}}$, such as the total number of segmented objects, are not affected by T value and only depend on the particular dataset. Larger values of ϵ (> 200) could have led to a slightly decrease in the number of segmented objects and a consequent increase in \bar{M}_A value; however, the influence of ϵ is minimal. On the contrary, it is evident the monotonically increasing width of A - and C -range for increasing T values. This clearly leads in parallel to a larger number of identified cells out of the total segmented objects. A suitable value of T can be chosen upon an evaluation of the tradeoff between false-positives and false-negatives (i.e. if too many cell clusters or cell debris are erroneously identified as isolated cells by MORPHEUS, T should be reduced; conversely, when too many true isolated cells are erroneously discarded, T can be increased). This task can be easily accomplished simply by scrolling *Img_#a - Segmentation* and *Img_#b - Detection* images in the *Cell* subfolder of the output directory.

Table SM3

T	Group	\bar{M}_A (pixel ²)	SE \bar{M} (pixel ²)	Area Range (pixel ²)		Circularity Range		Identified Cells	Detected Objects	%
				A_{low}	A_{high}	C_{low}	C_{high}			
1	EC+	3131.14	260.50	1435.32	6783.29	0.0601	0.1404	66	232	28.45
	EC-	3768.44	309.84	1729.30	8156.55	0.1036	0.2816	77	190	40.53
	PCC	1374.68	86.89	643.90	2923.14	0.3959	0.8733	330	591	55.84
2	EC+	3131.14	260.50	1305.07	7304.30	0.0562	0.1997	103	232	44.40
	EC-	3768.44	309.84	1574.38	8776.23	0.0962	0.3325	91	190	47.89
	PCC	1374.68	86.89	600.46	3096.91	0.3699	0.8777	352	591	59.56
3	EC+	3131.14	260.50	1174.82	7825.31	0.0523	0.2591	119	232	51.29
	EC-	3768.44	309.84	1419.46	9395.91	0.0888	0.3835	104	190	54.74
	PCC	1374.68	86.89	557.01	3270.68	0.3439	0.8821	374	591	63.28
4	EC+	3131.14	260.50	1044.56	8346.32	0.0485	0.3185	130	232	56.03
	EC-	3768.44	309.84	1264.54	10015.59	0.0813	0.4344	114	190	60.00
	PCC	1374.68	86.89	513.57	3444.45	0.3180	0.8865	414	591	70.05
5	EC+	3131.14	260.50	914.31	8867.33	0.0446	0.3778	133	232	57.33
	EC-	3768.44	309.84	1109.62	10635.26	0.0739	0.4853	118	190	62.11
	PCC	1374.68	86.89	470.13	3618.22	0.2920	0.8908	431	591	72.93
6	EC+	3131.14	260.50	784.06	9388.34	0.0407	0.4372	140	232	60.34
	EC-	3768.44	309.84	954.70	11254.94	0.0665	0.5363	128	190	67.37
	PCC	1374.68	86.89	426.69	3791.99	0.2660	0.8952	449	591	75.97

3.4. Orientation Analysis

As already stated, MORPHEUS uses the results of segmentation for two parallel tasks, namely morphometry (i.e. the evaluation of the shape descriptors for each cell) and orientation analysis. This latter task is actually performed only if the related check box in the starting dialog window of MORPHEUS was selected, and if *OrientationJ* plugin (version 2.0.2 or above) is present. In order to validate also this important feature, we used a fourth experimental condition represented

by human dermal fibroblast (HDF) cells cultured on a machined titanium dish. This particular substrate featured a number of parallel micrometric grooves serving as directional guides for cell growing (see *Cell Cultures and Image Acquisition* subsection). In particular, $m = 6$ fields of HDF cells fluorescently stained for cytoskeleton were acquired (for a total of $n = 96$ cells out of 131 detected objects) and, based on titanium machining, resulting images were rotated offline to make the grooves parallel to the x-axis (i.e. 0 degrees alignment). This procedure allowed verifying if—and to which extent—the same horizontal orientation had been induced on the cells plated over the titanium (Figure SM8A). Orientation analysis was conducted by MORPHEUS at both cellular and cytoskeleton scales: two representative HSB colormaps (see *Orientation Analysis* section) are shown in Figure SM8B and SM8C respectively, while coherency-weighted orientation histograms are showed in Figure SM8E in the form of heatmap matrices (i.e. *MasterMatrix_O*) and in Figure SM8D as total polar plots. At both cell and cytoskeleton scales, it is evident an enrichment of the 0°-orientation, thus confirming the steering power of the titanium substrate. Notice that, as expected, the orientation analysis performed at the cell level returned a more ‘discretized’ heatmap (Figure SM8E, left) if compared to the almost-continuous histograms coming from the cytoskeleton analysis (Figure SM8E, right). However, in order to have an actual quantification of such a steering effect, it turned out to be useful to consider the circular statistics for hypothesis testing and for the calculation of the main descriptors. Moreover, it is important to keep into account the axial nature of these particular data, which belong to a π -periodic domain (i.e. for a polarized cell, or a cytoskeleton element, 0° and 180° are indeed the same direction). By convention, in this case a domain from -90° (included) to +90° (excluded) was chosen. Even if the detailed mathematical formalism underlying the circular statistics cannot be addressed here, the most important descriptors of central tendency and dispersion are briefly recalled (see e.g. [10] for a good introduction to this wide yet still poorly explored branch of statistics). In general, when dealing with directional data, angles (measured in radians rather than in

degrees) are the random variables of interest. For each directional occurrence denoted by α_i (where $i = 1, 2, \dots, N$) a correspondent unit vector

$$r_i = (\cos \alpha_i, \sin \alpha_i)$$

can be defined. The *mean resultant vector*

$$\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i$$

is surely the most important circular statistic to be computed, since both the *mean angular direction* ($\bar{\alpha}$) and the *circular variance* (S) can be derived directly from it:

$$\bar{\alpha} = \arg(\bar{r})$$

$$S = 1 - R$$

where

$$R = \|\bar{r}\|$$

is the so-called *resultant vector length*, which is itself a measure of circular spread, since the closer it is to one, the more concentrated the data are around $\bar{\alpha}$. Notably, in contrast to the linear variance, S (and so R) is bounded in the interval $[0, 1]$. All these statistics can be retrieved from the total histogram returned by MORPHEUS as the final output of its directional analysis (*MasterMatrix_O*) and can be conveniently plotted using polar diagrams over the half-circle (Figure SM8D). In particular, HDF samples showed mean angular directions very close to 0° both at cell and cytoskeleton level ($\bar{\alpha} = 4.18^\circ$ and $\bar{\alpha} = 2.36^\circ$ respectively), whereas data exhibited a larger circular dispersion in the case of cytoskeleton analysis. However, for both the output, a Rayleigh test led to the rejection of the null hypothesis of uniform distribution over the circle. Moreover, a V-test assuming $\alpha_A = 0^\circ$ as mean direction for the alternative hypothesis, confirmed this unimodal deviation from uniformity (see Table SM4 for the related p -values and summary statistics).

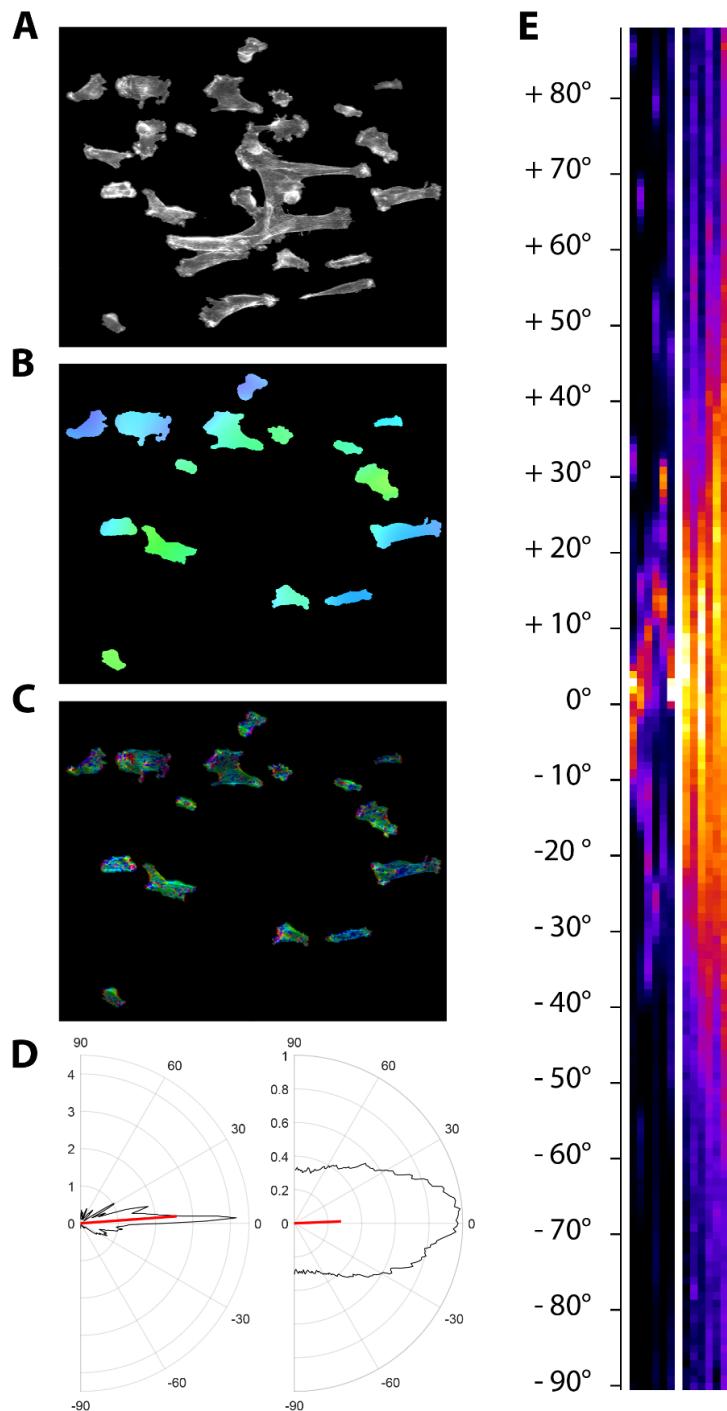


Figure SM8. Orientation analysis for HDF dataset. (A) A representative sample of Human Dermal Fibroblast (HDF) cells marked with phalloidin and cultured on a machined titanium dish. Images belonging to this dataset have been preliminarily rotated in order to align titanium grooves with x-axis. (B) Corresponding HSB colormap resulting from cell scale analysis. Hue encodes for the local dominant orientation, saturation encodes for coherency and brightness is based on the gray levels of the input image. (C) HSB colormap resulting from cytoskeleton scale analysis. In both B and C, notice the absence of the central cluster of cells, excluded by the preliminary MORPHEUS selection procedure. (D) Total coherency-weighted orientation histograms in the form of polar diagrams, for the analysis at both cell (left) and cytoskeleton (right) scale. Mean resultant vector is depicted in red. (E) MORPHEUS output MasterMatrix_O—containing orientation histograms of individual samples as columns—is represented as heatmap for the analysis at both cell (left) and cytoskeleton (right) resolution. Warmer regions are clearly located around 0° in both the submatrices (confirmed by significant Rayleigh and V-tests for circular uniformity).

Finally, as a negative control, we used the already described EC+ dataset, whose cells were not subjected to any orienting stimulus. As expected in these conditions, the two orientation heatmaps showed random patterns and circular statistics confirmed that no preferential direction was present neither at cell nor at cytoskeleton level (see Figure SM9 and Table SM4).

Notice that orientation data returned by MORPHEUS are binned by nature, since they come directly in the form of coherency-weighted orientation histograms with an angular resolution of 1° . However, MORPHEUS normalizes these histograms to the total number n of detected cells, thus providing an easy way to properly test them for circular uniformity. In other words, accounting for the sole isolated cells during the orientation analysis is not only a way to exclude those big cell clusters that are likely not to reflect the orientation of single cells, but it is also a necessary step for having an estimate of the total number of independent angular observations and then correctly conducting hypothesis tests. Also notice how this represents a quite conservative approach to the hypothesis testing, since n , as a measure of the number of independent angular observations, tends somehow to underestimate the real degrees of freedom of the samples, and consequently to overestimate p -values. This is especially true for cytoskeleton analysis, where the angular correlation among the filaments of same cell and the pixels thereof—even if still very high—is typically lower than the almost-unitary inherent correlation at cell scale.

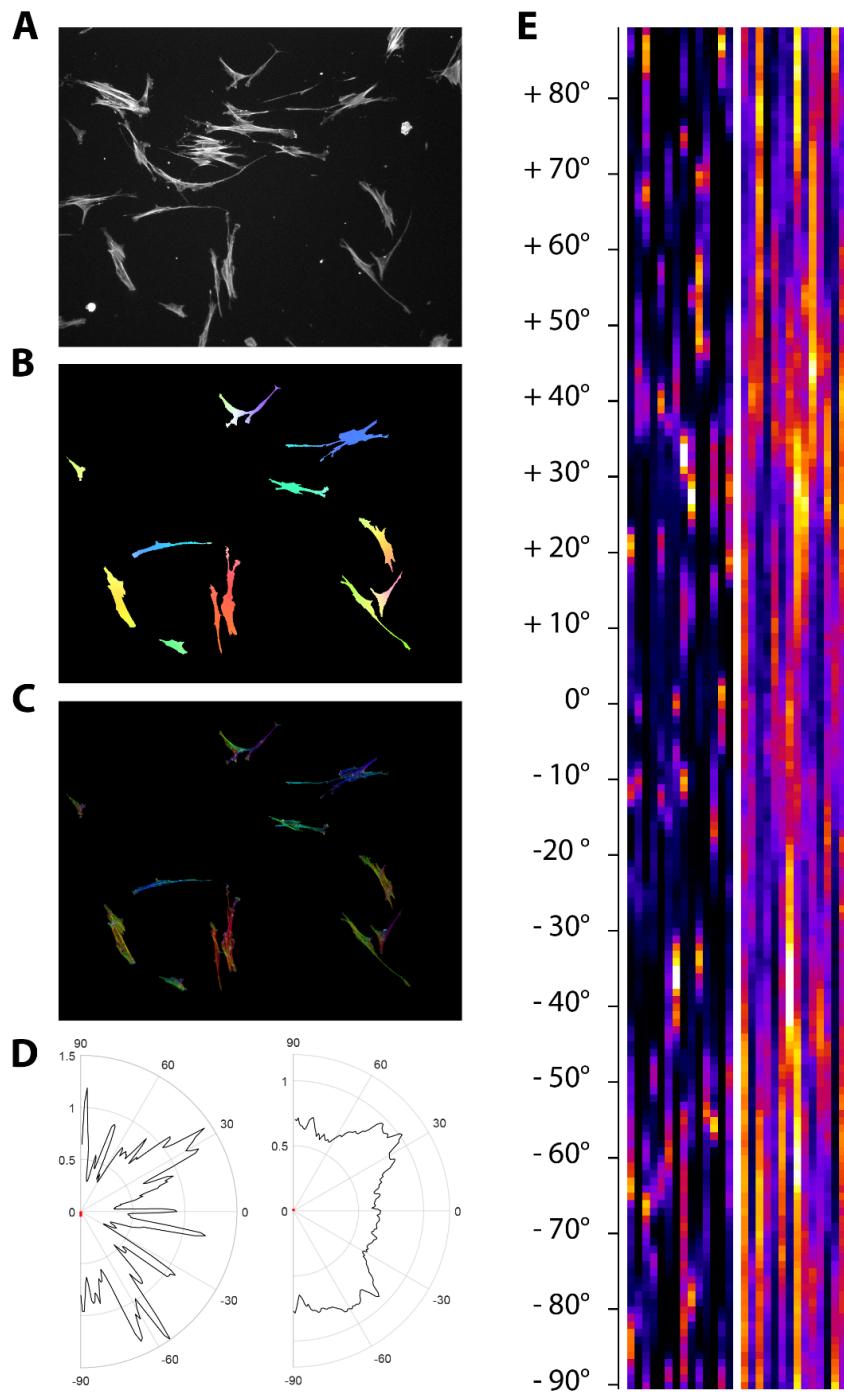


Figure SM9. Orientation analysis for EC+ sample. **(A)** A representative sample image of endothelial cells marked with phalloidin and cultured in the absence of any steering stimulus. **(B)** Corresponding HSB colormap resulting from cell scale analysis. Hue encodes for the local dominant orientation, saturation encodes for coherency and brightness is based on the gray levels of the input image. **(C)** HSB colormap resulting from cytoskeleton scale analysis. In both B and C, notice the absence of many cell clusters, excluded by the preliminary MORPHEUS selection procedure. **(D)** Total coherency-weighted orientation histograms in the form of polar diagrams, for the analysis at both cell (left) and cytoskeleton (right) scale. Notice the mean resultant vector collapsed in the origin, indicating the high value of angular variance and suggesting the absence of any preferential direction. **(E)** MORPHEUS output MasterMatrix_O—containing orientation histograms of individual samples as columns—is represented as heatmap for the analysis at both cell (left) and cytoskeleton (right) resolution. Both the submatrices exhibit a random pattern indicating no preferential direction (confirmed by the high p-values of Rayleigh and V-tests)

Table SM4

Dataset	HDF		EC+	
	Cell	Cytoskeleton	Cell	Cytoskeleton
Mean angular direction ($\bar{\alpha}$)	4.18°	2.36°	-89.92°	81.38°
Resultant vector length (R)	0.62	0.28	0.03	0.01
Circular Variance (S)	0.38	0.72	0.97	0.99
Rayleigh test p-value	$2.10 \cdot 10^{-18}$	$4.58 \cdot 10^{-4}$	0.8632	0.9718
V-test p-value	$< 10^{-16}$	$5.24 \cdot 10^{-5}$	0.7065	0.5905

4. Angular Resolution and Scale Dependency

As detailed in the *Orientation Analysis* section, MORPHEUS provides a quantitative assessment of the local dominant directions working over two distinct scales: cytoskeleton and whole-cell resolutions. Even if these two levels are typically highly correlated—in that cell orientation usually reflects cytoskeleton organization—it is worth noting that in principle the two procedures are completely independent, thus possibly leading to non-redundant results (i.e. uncorrelated coherency-weighted orientation histograms). Here we provide a demonstration of such an independence by taking advantage of 11 synthetic images, each one containing many elliptic objects featuring a ‘micro-structure’ orthogonally oriented respect to the ellipse major axis. Moreover, by means of the same set of images, also the 1° angular resolution of the algorithm will be evidenced. Images used for such a testing purpose are shown in Figure SM10. More in detail, each image is a collection of ellipses with the same dominant direction (major axis), but exhibiting slightly different size and variable eccentricity. In general, at whole-cell level, the greater the eccentricity, the greater the coherence. Conversely, at cytoskeleton scale, the lined pattern within each object was designed to have the same dominant direction of ellipse minor axis (90°-shifted respect to whole-cell orientation) and always provides high coherency values. A detailed description of each sample image can be found in the caption of Figure SM10, while Figure SM11 and Figure SM12 contain the HSB colormaps resulting from MORPHEUS orientation analysis at whole-cell and cytoskeleton scale respectively. Finally, Figure SM13 shows the heatmap representing the 11 coherency-weighted orientation histograms, evaluated for both the spatial scales (*MasterMatrix_O*). Since each single sample image contains just one particular dominant direction—both at whole-cell and cytoskeleton level—histograms appear as 1°-narrow peaks, corresponding to single pixels in the heatmap. By comparing the original images—and the information therein—with the HSB colormaps and the heat-mapped histograms it is possible to test for OrientationJ and MORPHEUS accuracy (correspondence between theoretical and measured

direction), precision (spread of histogram peaks), angular resolution (capability of discriminating two close dominant directions) and sensitivity (capability of detecting weak signals coming from low-coherency objects or representing minority sub-populations). Most importantly, by comparing left and right side of Figure SM13, it is possible to verify the complete independence between the two different levels of analysis accomplished by MORPHEUS.

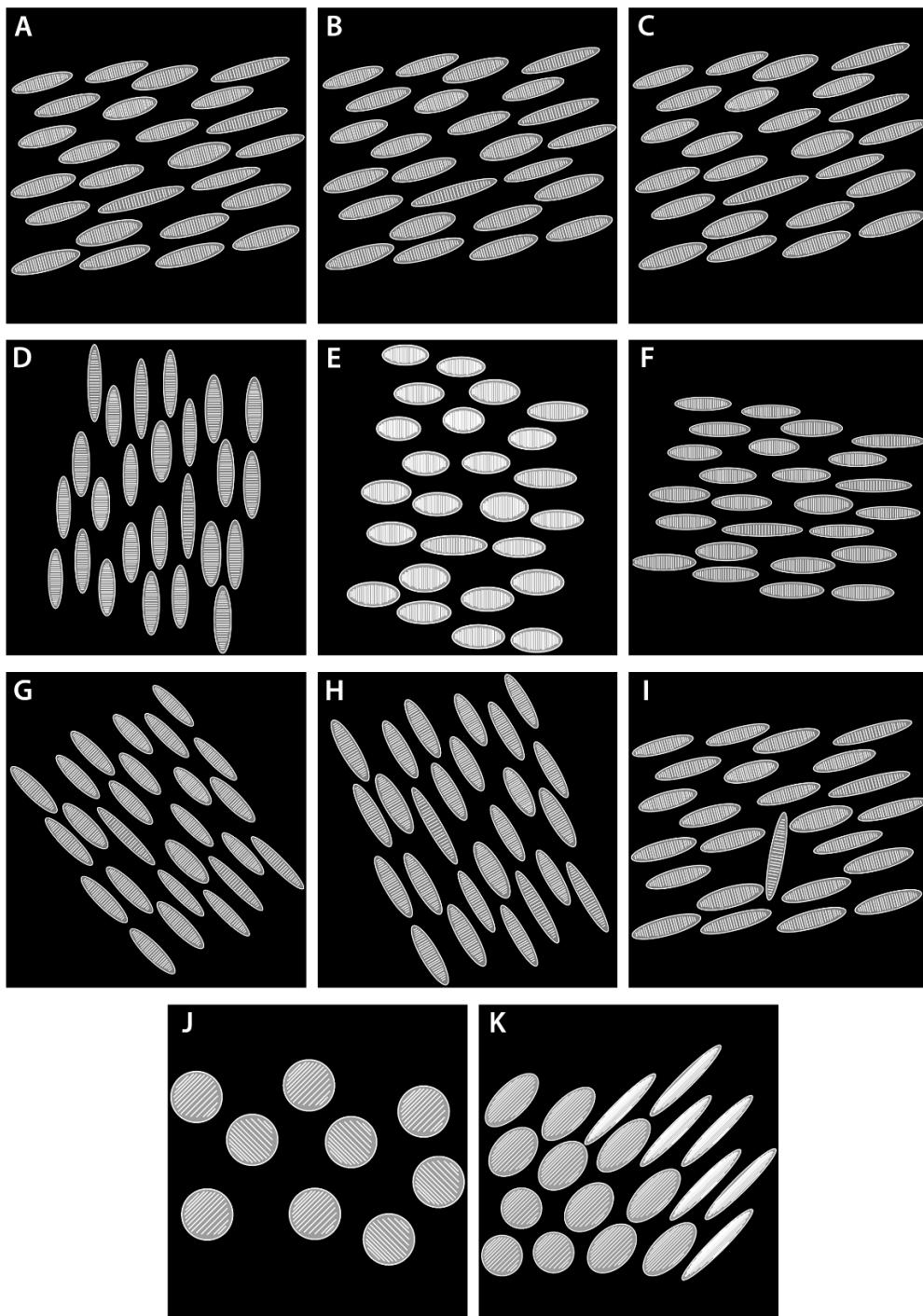


Figure SM10. Catalogue of the 11 synthetic images used for validating the performance of MORPHEUS in the context of directional analysis. Each image contains many ellipses with a common principal direction ϑ that is the orientation of their major axis. In addition, each ellipse features an internal pattern of lines with an independent orientation δ . For this test, δ has been conventionally chosen to be that of ellipse minor axis ($\delta = \vartheta \pm 90^\circ$). **(A)** $\vartheta = +15^\circ$ and $\delta = -75^\circ$. **(B)** $\vartheta = +16^\circ$ and $\delta = -74^\circ$. **(C)** $\vartheta = +17^\circ$ and $\delta = -73^\circ$. **(D)** $\vartheta = -90^\circ$ and $\delta = 0^\circ$. **(E)** $\vartheta = 0^\circ$ and $\delta = -90^\circ$ with a lower mean eccentricity. **(F)** $\vartheta = 0^\circ$ and $\delta = -90^\circ$ with a higher mean eccentricity. **(G)** $\vartheta = -45^\circ$ and $\delta = +45^\circ$. **(H)** $\vartheta = -60^\circ$ and $\delta = +30^\circ$. **(I)** $\vartheta = +15^\circ$ and $\delta = -75^\circ$ with an anomalous object having $\vartheta = +80^\circ$ and $\delta = -10^\circ$. **(J)** ϑ = undetermined (circles with eccentricity = 0) and $\delta = \pm 45^\circ$. **(K)** $\vartheta = +45^\circ$ with a variable eccentricity (progressively orientated objects) and $\delta = +45^\circ$ (exceptionally equal to ϑ).

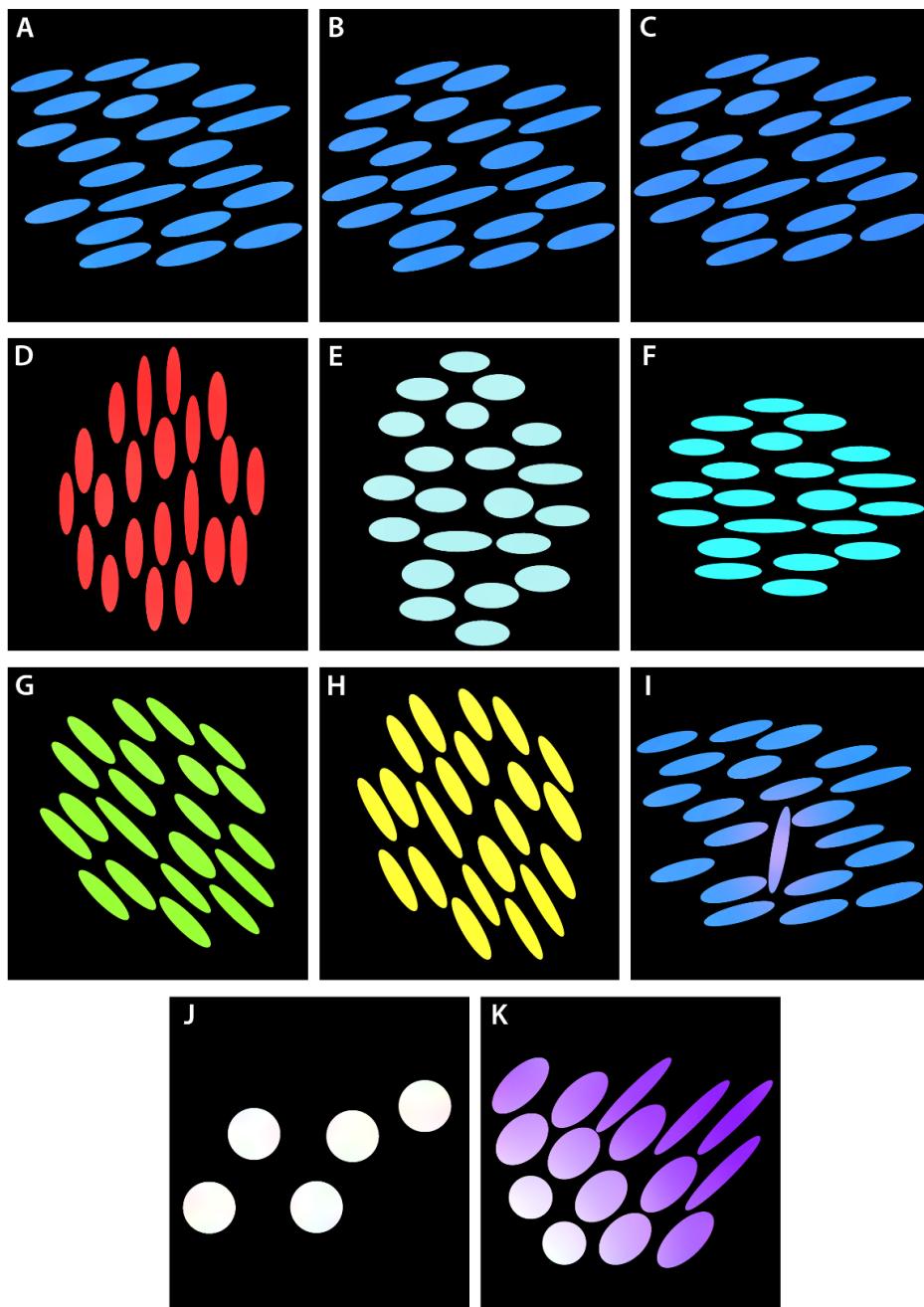


Figure SM11. HSB colormaps resulting from MORPHEUS orientation analysis at whole-cell scale. In order to perform the analysis with a such spatial resolution, the radius of the sliding Gaussian window was set to $\sqrt{\bar{M}_A}/\pi$, where \bar{M}_A was the characteristic area of the ‘single cell’, namely the expected value of sample median area. Colormaps have been labeled according to the original samples of Figure SM10. **(A,B,C)** The difference of 1° among the major axis orientation of the first three samples, even if not detectable by eye, is accurately detected by the algorithm (see first three columns of the left heatmap in Figure SM13). **(D,E,F,G,H)** Different orientations are coded by different hue values, while saturation represents the local coherency of the objects (eccentricity of the ellipses). As an example, notice the different saturation of E and F, though sharing the same hue/principal direction ($\vartheta = 0^\circ$). **(I)** Even if with some interference from neighboring objects, the algorithm exhibited a high sensitivity proving able to detect the contribution of a minority subpopulation made of just one object. **(J,K)** Circles can be thought as zero-eccentricity degenerate ellipses, so they have a null coherency and appear completely desaturated (i.e. white) in the HSB colormap.

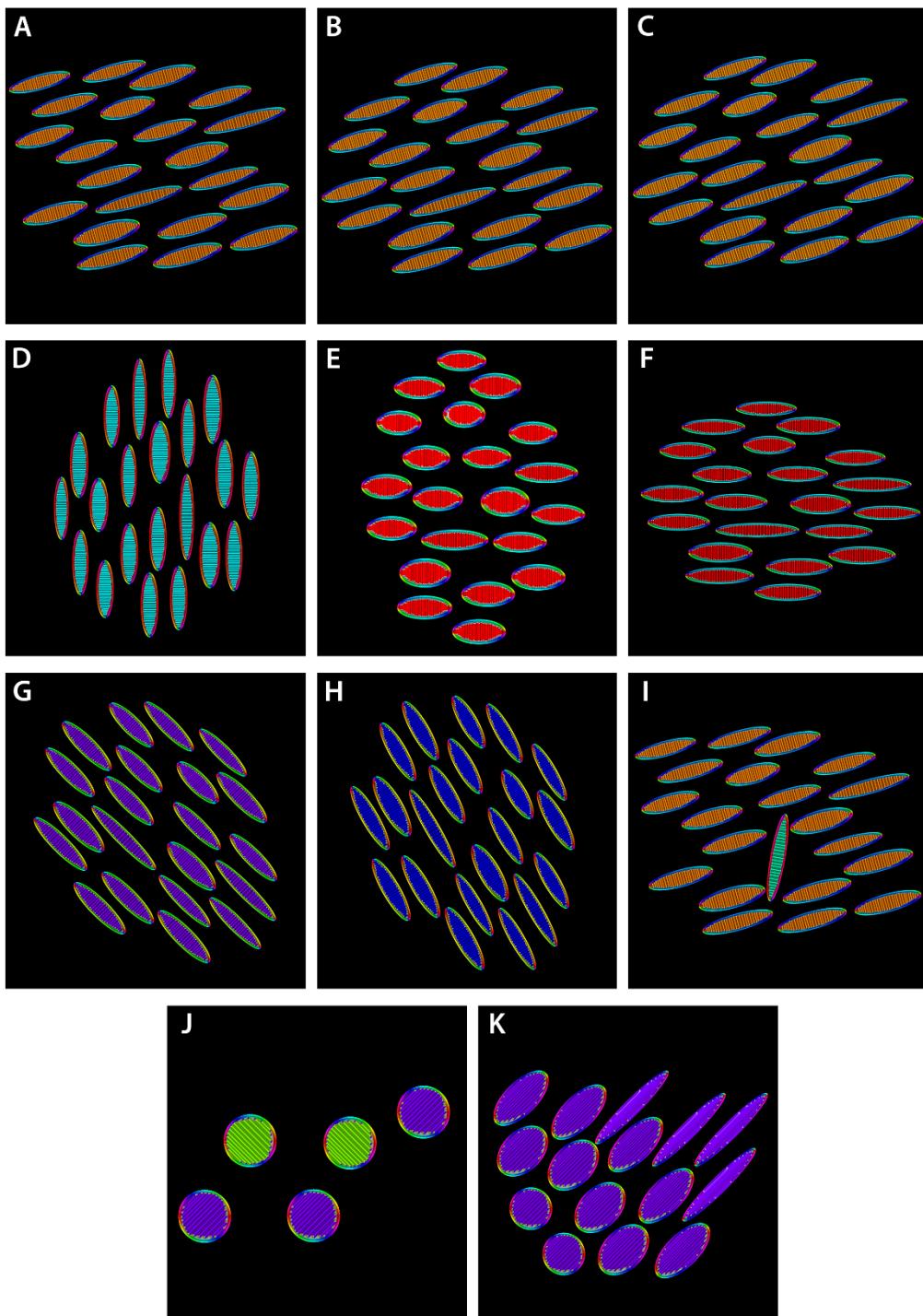


Figure SM12. HSB colormaps resulting from MORPHEUS orientation analysis at cytoskeleton scale. In order to perform the analysis with a such spatial resolution, the radius of the sliding Gaussian window was set to its minimum value, namely 1 pixel. Colormaps have been labeled according to the original samples of Figure SM10. Regardless of the coherency of the entire objects, the internal lined structure of the ellipses always provided high coherency values when analyzed at cytoskeleton scale. To this purpose, notice in particular subpanels J and K and compare them with the corresponding ones in Figure SM11.

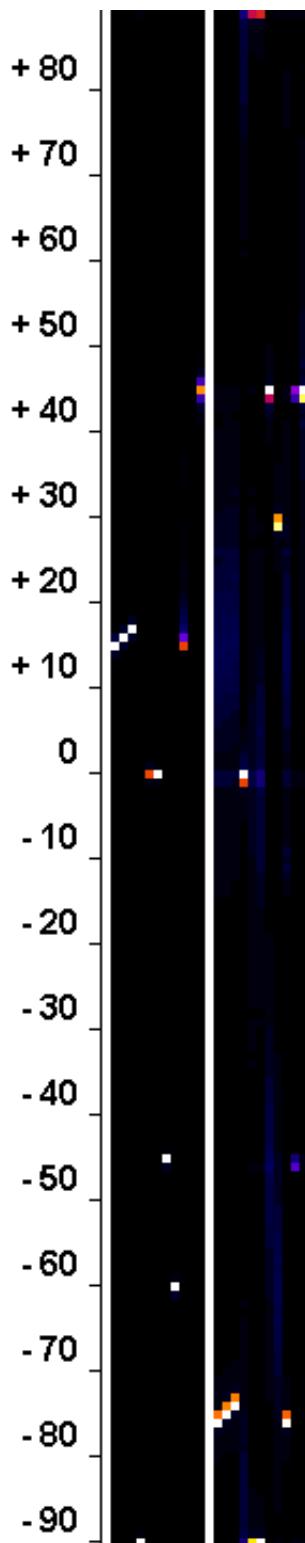


Figure SM13. Heatmap of the coherency-weighted orientation histograms. The principal direction characteristic of each sample image is here represented by a single pixel (narrow angular distribution) both at whole-cell and at cytoskeleton scale (left and right panel, respectively). The shift of 90° between the two submatrices reflects the particular structure of the synthetic objects used to test the independence between the two levels of analysis (ellipses with an internal pattern made of lines parallel to the minor axis). By comparing these heatmaps with the original images (Figure SM10) and the related colormaps (Figure SM11-12), the user can check the reliability of the orientation analysis performed by MORPHEUS.

5. References

- [1] I. Arganda-Carreras and P. Andrey, "Designing Image Analysis Pipelines in Light Microscopy: A Rational Approach.", *Methods Mol. Biol.*, vol. 1563, pp. 185–207, 2017.
- [2] S. Sternberg, "Biomedical Image Processing," *Computer (Long. Beach. Calif.)*, vol. 16, no. 1, pp. 22–34, Jan. 1983.
- [3] C. H. Li and P. K. S. Tam, "An iterative algorithm for minimum cross entropy thresholding," *Pattern Recognit. Lett.*, vol. 19, no. 8, pp. 771–776, Jun. 1998.
- [4] Z. Püspöki, M. Storath, D. Sage, and M. Unser, "Transforms and Operators for Directional Bioimage Analysis: A Survey," in *Focus on Bio-image informatics*, Springer, Ed. 2016, pp. 69–93.
- [5] R. Rezakhanlha *et al.*, "Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy," *Biomech. Model. Mechanobiol.*, vol. 11, no. 3–4, pp. 461–73, Mar. 2012.
- [6] E. Fonck *et al.*, "Effect of Aging on Elastin Functionality in Human Cerebral Arteries," *Stroke*, vol. 40, no. 7, pp. 2552–2556, Jul. 2009.
- [7] E. Batschelet, *Circular Statistics in Biology Mathematics in biology Quantitative Studies in Social Relations*. 1981.
- [8] S. Jammalamadaka, *Topics in Circular Statistics*. 2001.
- [9] J. Zar, *Biostatistical Analysis*. 1999.
- [10] N. I. Fisher, *Statistical Analysis of Circular Data*. 1995.
- [11] G. K. Sandve, A. Nekrutenko, J. Taylor, and E. Hovig, "Ten Simple Rules for Reproducible Computational," vol. 9, no. 10, pp. 1–4, 2013.
- [12] N. Basilico *et al.*, "Dextran-shelled oxygen-loaded nanodroplets reestablish a normoxia-like pro-angiogenic phenotype and behavior in hypoxic human dermal microvascular endothelium," *Toxicol. Appl. Pharmacol.*, vol. 288, no. 3, pp. 330–338, Nov. 2015.
- [13] S. Petrillo *et al.*, "Heme accumulation in endothelial cells impairs angiogenesis by triggering paraptosis," *Cell Death Differ.*, pp. 1–16, 2018.
- [14] T. Genova *et al.*, "TRPM8 inhibits endothelial cell migration via a non-channel function by trapping the small GTPase Rap1," *J. Cell Biol.*, vol. 216, no. 7, pp. 2107–2130, Jul. 2017.
- [15] F. Mussano *et al.*, "Title: In vitro characterization of two different atmospheric plasma jet chemical functionalizations of titanium surfaces," *Appl. Surf. Sci.*, vol. 409, pp. 314–324, 2017.
- [16] L. Canullo, T. Genova, P. Mandracci, F. Mussano, R. Abundo, and J. Fiorellini, "Morphometric Changes Induced by Cold Argon Plasma Treatment on Osteoblasts Grown on Different Dental Implant Surfaces," *Int. J. Periodontics Restorative Dent.*, vol. 37, no. 4, pp. 541–548, Jul. 2017.
- [17] L. Canullo, T. Genova, H.-L. Wang, S. Carossa, and F. Mussano, "Plasma of Argon Increases Cell Attachment and Bacterial Decontamination on Different Implant Surfaces," *Int. J. Oral Maxillofac. Implants*, vol. 32, no. 6, pp. 1315–1323, Nov. 2017.
- [18] A. Patelli *et al.*, "Nanoroughness, Surface Chemistry, and Drug Delivery Control by Atmospheric Plasma Jet on Implantable Devices," *ACS Appl. Mater. Interfaces*, vol. 10, no. 46, pp. 39512–39523, Nov. 2018.
- [19] F. Mussano *et al.*, "Hydrogenated amorphous silicon coatings may modulate gingival cell response," *Appl. Surf. Sci.*, vol. 436, pp. 603–612, 2018.

MORPHEUS – Supplementary Materials

- [20] F. Mussano, T. Genova, F. Serra, M. Carossa, L. Munaron, and S. Carossa, “Nano-Pore Size of Alumina Affects Osteoblastic Response,” *Int. J. Mol. Sci.*, vol. 19, no. 2, p. 528, Feb. 2018.
- [21] F. Mussano *et al.*, “Early Response of Fibroblasts and Epithelial Cells to Pink-Shaded Anodized Dental Implant Abutments: An In Vitro Study,” *Int. J. Oral Maxillofac. Implants*, vol. 33, no. 3, pp. 571–579, May 2018.
- [22] A. V. Frane, “Power and Type I Error Control for Univariate Comparisons in Multivariate Two-Group Designs,” *Multivariate Behav. Res.*, vol. 50, no. 2, pp. 233–247, Mar. 2015.