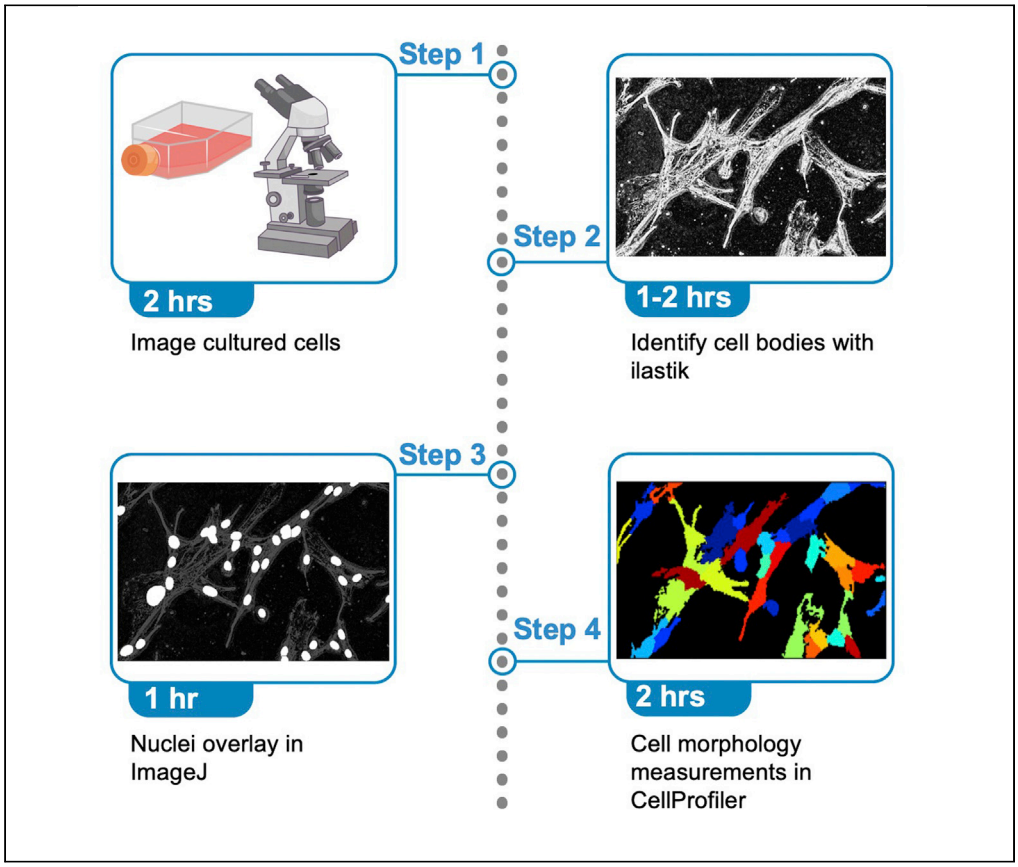
STAR Protocols

CellPress
OPEN ACCESS

Protocol

An open access, machine learning pipeline for high-throughput quantification of cell morphology



Cell morphology is influenced by many factors and can be used as a phenotypic marker. Here we describe a machine-learning-based protocol for high-throughput morphological measurement of human fibroblasts using a standard fluorescence microscope and the pre-existing, open access software ilastik for cell body identification, ImageJ/Fiji for image overlay, and CellProfiler for morphological quantification. Because this protocol overlays nuclei with their cell bodies and relies on coloration differences, it can be broadly applied to other cell types beyond fibroblasts.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Emma M. Welter, Oksana Kosyk, Anthony S. Zannas

emwelter21@gmail.com (E.M.W.)
anthony_zannas@med.unc.edu (A.S.Z.)

Highlights

Machine learning pipeline for quantifying cell morphology

Uses standard fluorescence microscope and open access software

Implemented in IMR-90 fibroblasts but adaptable to other cell types

Welter et al., STAR Protocols 4, 101947
March 17, 2023
https://doi.org/10.1016/j.xpro.2022.101947

STAR Protocols

CellPress
OPEN ACCESS

Protocol

An open access, machine learning pipeline for high-throughput quantification of cell morphology

Emma M. Welter,[1,3,*] Oksana Kosyk,[1] and Anthony S. Zannas[1,2,4,*]

[1]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[2]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[3]Technical contact

[4]Lead contact

*Correspondence: emwelter21@gmail.com (E.M.W.), anthony_zannas@med.unc.edu (A.S.Z.)
https://doi.org/10.1016/j.xpro.2022.101947

SUMMARY

Cell morphology is influenced by many factors and can be used as a phenotypic marker. Here we describe a machine-learning-based protocol for high-throughput morphological measurement of human fibroblasts using a standard fluorescence microscope and the pre-existing, open access software ilastik for cell body identification, ImageJ/Fiji for image overlay, and CellProfiler for morphological quantification. Because this protocol overlays nuclei with their cell bodies and relies on coloration differences, it can be broadly applied to other cell types beyond fibroblasts.

For details on the use and execution of this protocol, please also refer to Leung et al. (2022).[1]

BEFORE YOU BEGIN

The present protocol enables the quantification of cell morphology, and it has been implemented in human fetal lung fibroblasts (IMR-90).[1] Detailed here is the procedure with the software parameters set for this cell type and the magnification settings chosen for our images. When adapting this pipeline for use in other cell types or contexts, all parameters must be checked by visual inspection and the conversion factor between pixels and micrometers must be recalculated. A schematic of the processing steps can be seen in Figure 1. This protocol can be adapted for other cell types, as it primarily relies on distinguishing cell body from image background. While the presented work focuses on cell surface area, perimeter, circularity, pseudopodia number, and total pseudopodia length, additional morphological measurements can be derived with the used software.

The presented pipeline utilizes three open access software: ilastik version 1.3.3post3,[2] ImageJ/Fiji version 1.53e,[3] and CellProfiler version 4.1.3.[4] There are many available resources on how to operate these programs. The settings described here for each of these programs are what was found to work best for our images and may not apply to all microscopes or cell types. Changes may have to be made when using later versions of each software.

Institutional permissions (if applicable)

No institutional permissions are applicable.

Cell culture

Timing: hours to months (depending on treatment of interest)

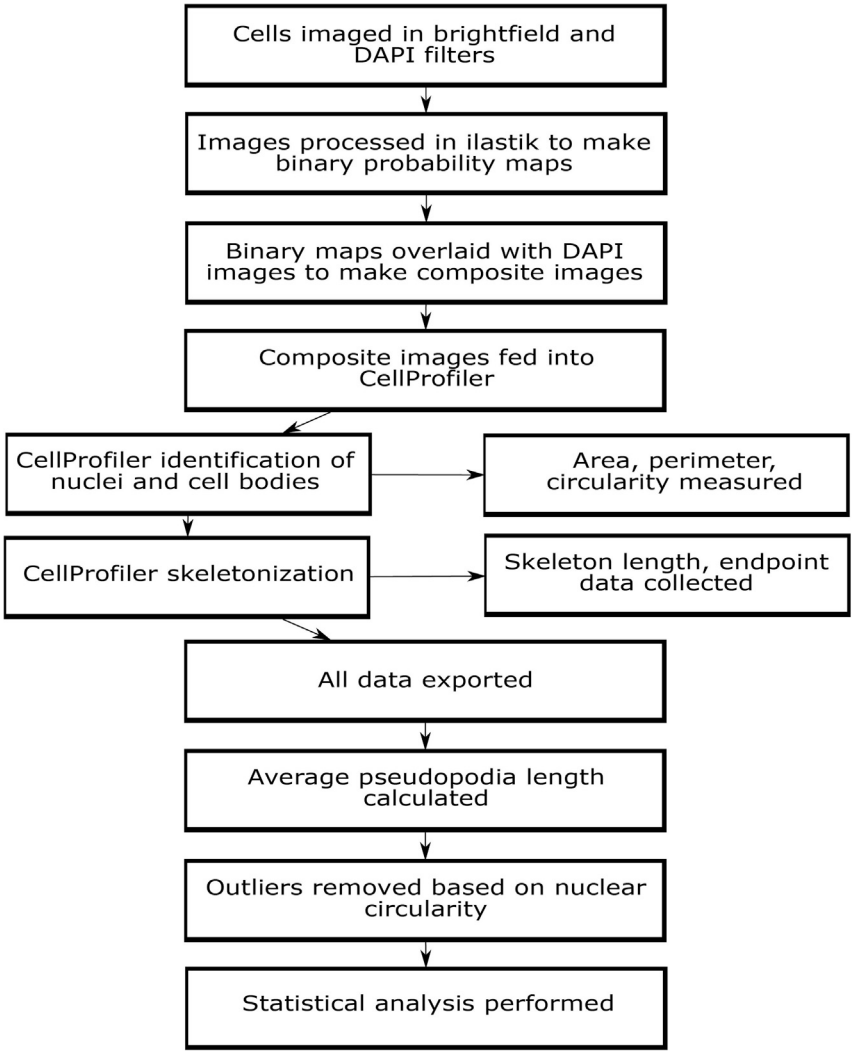


Figure 1. Schematic overview of the pipeline's processing and quantification steps

In this protocol, we used a human fetal lung fibroblast cell-line (IMR-90 cells). The cells were obtained from the Coriell Institute and stored at vapor phase of liquid nitrogen prior to cell culture. Upon recovery from cryopreservation, cells were maintained in 1× DMEM medium supplemented according to the Complete Growth Media table in a humified 37°C incubator and 5% CO2. Cells were subcultivated every 3–4 days, detached using 0.05% Trypsin-EDTA, and seeded into a 6-well plate for morphological analysis. The cultured media, cryopreservation media, solutions and reagents were prepared by mixing the reagents as listed in materials and equipment below.

1. Obtain or thaw cells for experiment.
2. Culture Cells according to desired treatment.
   a. Warm culture media, PBS, and Trypsin in a water bath.
   b. Sterilize a cell culture hood with 70% Ethanol and have all required supplies on hand.
   c. Bring a cell flask from an incubator to a hood, making sure to handle flask with care.

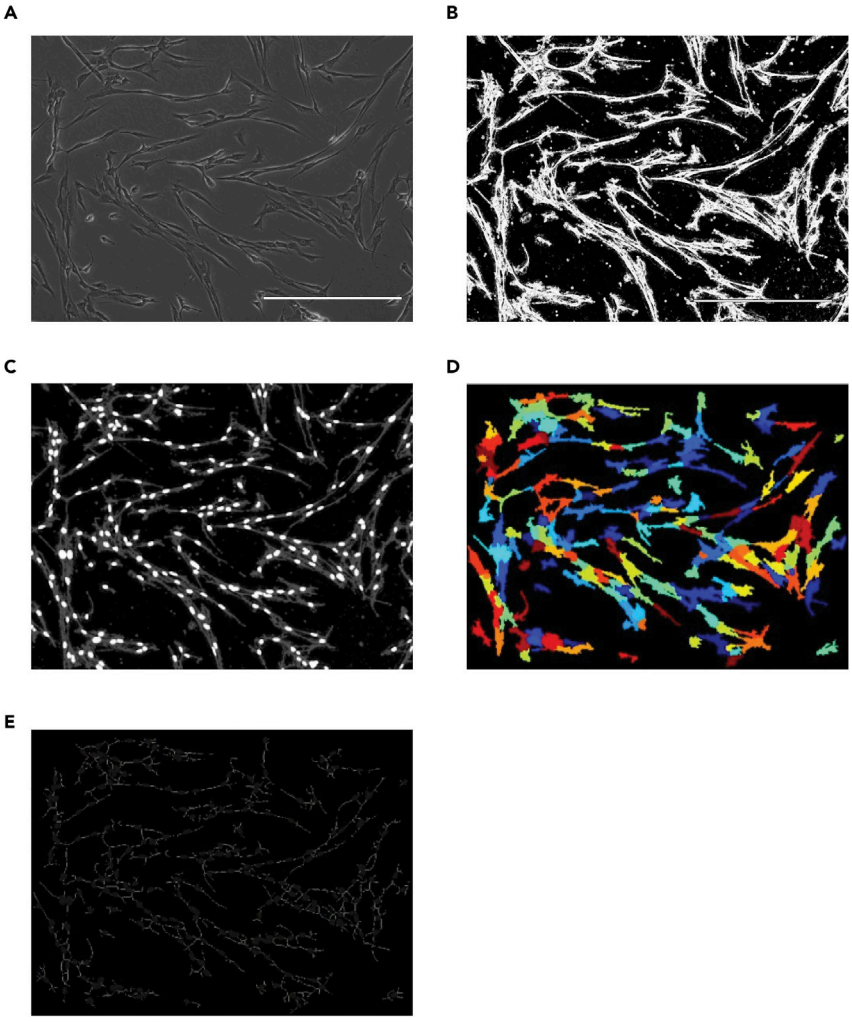


Figure 2. Processing stages of the morphology quantification pipeline, using representative images from early passage IMR-90 cells

(A) Original brightfield image taken with EVOS FL microscope (scalebar is equivalent to 400 μm).
(B) ilastik identification of cell bodies (scalebar is equivalent to 400 μm).
(C) Composite image created from the overlay of ilastik and DAPI images.
(D) CellProfiler identification of cell bodies.
(E) CellProfiler skeletonization for pseudopodia quantification.

d. Remove the media from the flask using a sterile pipette.
e. Rinse the cells with 1× PBS.

Note: Gently semi-rotate the flask to ensure that the solution is fully covering the cells.

f. Remove 1× PBS and discard it into a waste container.
g. Add 0.05% Trypsin-EDTA or another appropriate cell dissociation solution to the flask.
h. Gently semi-rotate the plate to ensure that the solution is covering the surface.
i. Incubate the cells at 37°C for 5–7 min.
j. Inspect the cells under microscope to confirm that they have detached from the bottom of the flask.

Note: Once detached, the cells will become round and float in the media if flask is tilted.

k. Inactivate Trypsin by adding two volumes (2× volume used in step 2g) of cell culture media supplemented with fetal bovine serum.
l. Transfer the cells into 50 mL conical tube and centrifuge at 100 ×G for 5 min to pellet the cells.

Note: Upon completion of the centrifugation, there should be a well-formed pellet.

m. Remove the media with a sterile pipette and discard it, ensuring not to disturb the cell pellet.
n. Resuspend the cells in 2 mL of complete DMEM media by gently pipetting cells up and down until there is a homogeneous cell solution.
o. Remove a small (10 μL recommended) aliquot of the cell solution for cell counting.
p. Use the Trypan Blue assay with the Countess II FL automated cell counter to count the cells and determine the ratio of live/dead cells.

Note: Combine 10 μL aliquot of cell solution to 10 μL Tryphan Blue and load into cell counting chamber.

q. Based on cell counting, determine the number of cells and volume of media to be used for optimal seeding density of $0.2*10^6$–$0.25*10^6$ cells per well (recommended if measuring human fibroblasts) on a 6-well culture plate.

Note: As this protocol utilizes two-dimensional images of three-dimensional cells, maintaining a consistent confluency and seeding density controls for the amount of overlap between cells. We have found the optimal final confluency to be 40%–50% for imaging.

r. Mix the cells with fresh media gently and dispense 3 mL per well of the solution to a new 6-well plate.
s. Semi-tilt to evenly distribute the cells of each well.
t. Place the plate into the incubator and leave it undisturbed for 24 h.

Optional: Synchronize cell cycle stage using Aphidicolin.

u. Remove an aliquot of aphidicolin from freezer and thaw on ice.

Note: In our experience, subjecting aphidicolin readymade solution to multiple freeze/thaw cycles is acceptable, though it should still be avoided if possible. However, aphidicolin is light sensitive and must be protected from direct light exposure during storage.

v. Bring previously seeded cells in 6-well plates from the incubator to the cell culture hood.
w. Inspect cells under microscope, ensuring that they have attached to the bottom of the well and uniformly cover the surface with a confluency between 40%–50%.
x. Add 1 μL of aphidicolin per 1 mL of media directly into each well.[5,6]
y. Mix gently by semi-tilting the plate and return plate to incubator.
z. Incubate plate for 12–16 h to ensure synchronization of cell cycle.

3. Fix cells for imaging.
   a. Chill methanol at −20°C for at least one hour prior to use.
   b. Remove plates from incubator, and remove media from each well.
   c. Wash cells twice with 1× PBS.
   d. Remove 1× PBS and immediately add chilled methanol to each well by dispensing it gently on the side of the well.

Note: Avoid pipetting Methanol directly on the cells as it may result in cell detachment.

e. Incubate the cells at −20°C for 10 min.
f. Check the cells under the microscope to ensure proper fixation.

Note: The majority of cells should remain intact with clear morphology (Figure 2A). See Troubleshooting problem 1.

g. Remove methanol from the well and wash again with 1× PBS.
h. Fill the well with 1× PBS following the wash.
i. Prepare DAPI solution according to manufacturer instructions.

Note: The DAPI solution used here is provided at a 1 mg/mL concentration by the manufacturer and is diluted 1:1000 with 1× PBS for a 1 μg/mL final working concentration. The stock DAPI solution should be stored at −20°C and protected from light. It is recommended to aliquot the readymade solution and avoid freeze/thaw cycles. The prepared DAPI solution can be stored at 4°C for six months.

j. Incubate cells with DAPI for 2 min to stain the nuclei, with the plates shielded from light.

Note: One drop (~20–30 μL) of the diluted DAPI solution from the provided bottle is sufficient for staining at the recommended cell density of $0.2*10^6$–$0.25*10^6$ cells per well.

Optional: The above step 2u-z of the cell culture section is optional. Aphidicolin is a compound that synchronizes the cell cycle at the G1/S boundary.[5,6] Synchronizing the cell cycle reduces noise from the subsequent morphological measurements. This step is important when experimental conditions can influence the proliferation rate of cells (for example, comparisons between different cell passages).

Imaging

Timing: 2 h

This section details the imaging of the plate for further analysis. This is best done immediately following fixation of the cells. This imaging study utilizes an EVOS FL microscope, which is non-automated and possesses both DAPI and brightfield filters. Other microscope types (including automated microscopes) can be used for this procedure, as the main instrumental requirement is the ability to take both brightfield and DAPI-stained images of the same field.

4. Place the 6-well plate with fixed and stained cells on microscope stage.

Note: It is vital to use a microscope that is capable of imaging with both the brightfield and DAPI filters without moving the stage/visual field of the microscope.

5. Focus the microscope and choose an image field where the cells are well represented.

Note: This can be done by choosing fields with intact cells that are similar to the average phenotype of the cells present in the well. An example can be seen in Figure 2A.

6. Capture and save the brightfield image.

Note: Make sure to keep the lighting in the brightfield images as consistent as possible. This can often be controlled by using the same amount of light to illuminate the images. For the

EVOS FL microscope, a range of 43%–47% light was appropriate. See Troubleshooting problem 2.

7. Change the filter on the microscope to one that highlights the DAPI stained nuclei and image and save the exact same field used for the brightfield image.

   Note: The EVOS FL microscope utilizes a Class 3B ultraviolet LED light cube with a 360 nm excitation and 447 nm emission for DAPI imaging.

8. Repeat process for all wells in the plate (imaging 6–8 regions per well is recommended).

   Note: When using a non-automated microscope, it is best to use a magnification that is high enough to allow imaging multiple regions within the same well without overlap. For the EVOS FL microscope, we utilized a 10× objective. We suggest moving the microscope stage between images to avoid imaging overlapping regions.

9. Load saved images onto the computer.
10. Open the ImageJ/Fiji program, and convert the image type of the brightfield images to RGB color. (Image > Type > RGB Color).

    Note: This may not be necessary depending on the format of microscope-exported images.

    ⚠ CRITICAL: When imaging the brightfield and DAPI images (steps 6–9), ensure that the stage is not moved. These two images will be overlaid later, so they need to be of the exact same field.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
DMEM, high glucose, no glutamine, no phenol red	Gibco	Cat # 31053028
MEM non-essential amino acids solution (100×)	Gibco	Cat # 11140050
L-Glutamine (200 mM)	Gibco	Cat # 25-030-081
Sodium pyruvate (100 mM)	Gibco	Cat # 11360070
Antibiotic-antimycotic (100×)	Gibco	Cat # 15240062
Avantor Seradigm Premium Grade Fetal Bovine Serum (FBS)	VWR	Cat # 97068-085
Dimethyl sulfoxide	Millipore Sigma	Cat # D2650
Aphidicolin Ready Made Solution	Millipore Sigma	Cat # A4487
Methanol	Millipore Sigma	Cat # 34860
DAPI Ready Made Solution	Millipore Sigma	Cat # MBD0015
Trypsin-EDTA	Millipore Sigma	Cat # T4174
Experimental models: Cell lines		
IMR-90 human female fetal lung fibroblasts	Coriell Institute for Medical Research	Cat # I90-10
Software and algorithms		
Ilastik version 1.3.3post3	Berg et al., 2019	https://www.ilastik.org/
CellProfiler version 4.1.3	Carpenter et al., 2006	https://www.cellprofiler.org/
ImageJ version 1.53e	Rasband, 2011	https://imagej.nih.gov/ij/
Other		
EVOS FL microscope	EVOS Cell Imaging Systems	Cat # AMF4300R
Countess II FL automated cell counter	Invitrogen	Cat # AMQAF1000

MATERIALS AND EQUIPMENT

Complete growth media		
Reagent	Final concentration	Amount
Non-Essential Amino Acids 100×	1×	5 mL
L-Glutamine 100×	1×	5 mL
Sodium Pyruvate 100×	1×	5 mL
Antibiotic-Antimycotic 100×	1×	5 mL
Fetal Bovine Serum	0.15×	75 mL
DMEM Medium 1×	0.81×	405 mL
Total		**500 mL**
Cell culture media can be stored up to 4 weeks at 4°C and must be brought to 37°C prior to use.		

Cryopreservation media		
Reagent	Final concentration	Amount
Non-Essential Amino Acids 100×	1×	1 mL
L-Glutamine 100×	1×	1 mL
Sodium Pyruvate 100×	1×	1 mL
Fetal Bovine Serum	0.2×	20 mL
DMEM Medium 1×	0.67×	67 mL
DMSO	0.1×	10 mL
Total		**100 mL**
Cryopreservation medium must be made fresh prior use and chilled to 4°C on ice.		

⚠ CRITICAL: Methanol is highly flammable; acutely toxic orally, dermally, and if inhaled; and has organ-specific toxicity. It should be handled with appropriate gear (protective gloves, clothing, eye protection and face protection). DAPI may cause skin irritation and fumes should be avoided, protective gloves should be used. DMSO is a flammable liquid that should be handled appropriately and with proper protective gear. Sodium pyruvate can cause serious eye damage and possible skin reaction and should be handled with face protection in addition to standard protective gear in a well-ventilated area. Antibiotic-Antimycotic may damage fertility or cause respiratory issues and should be handled with appropriate protective gear in a well-ventilated area. Aphidicolin readymade solution is flammable and should be handled with appropriate protective gear (protective gloves, eye protection, face protection).

STEP-BY-STEP METHOD DETAILS

Binary segmentation in ilastik

Timing: 1–2 h

This section is the machine learning component of the pipeline. Here, the program must be taught what each class of pixel represents in order to segment the images accordingly. This results in binary images with the cell body distinct from the background (Figure 2B). Once training is completed for your dataset, it will likely not need to be repeated. All major steps listed here can theoretically be paused at any point, so long as you keep record of where you left off.

1. Open the previously trained ilastik Pixel Classification file and set the desired export location of the resulting Prediction Maps (binary representations of cell body and background). This is done by choosing Prediction Export > Choose Export Image Settings > Select Directory.

Note: The data should be converted to unsigned 16-bit, renormalized from 0.00, 1.00 to 0, 65535, the axis order should be transposed to axis order cyx, and the output file format should be tif sequence, though settings may vary depending on microscope type. The File Pattern should be "nickname_result_type_slice_index.tif". If there is no previously trained ilastik model or its training needs to be altered, see the following sub-steps.

a. For preparing a completely new model, open a new Pixel Classification module and choose brightfield images that are representative of the dataset. Input these into the Input Data section.

Note: We recommend including at least one representative image for each condition, with each condition equally represented.

i. Select the sigma value features under Feature Selection,[7] which enable the computer to account for different image characteristics, such as color, edge, and texture, when making decisions for classifying cell shapes.

Note: Start with the lower end of the sigma values and change as necessary later on. Our most recent iteration of this pipeline has found the sigma values of 0.70 for color/intensity, 1.00 for edge, and 1.60 for texture to be effective. See Troubleshooting problem 3.

ii. Add three classes of labels under Training for the Cell Body, Background, and Boundary of the image.[8]

Note: The boundary label refers to the region between the cell body and background, which is often a region of high lighting variation in brightfield images. We suggest annotating a minimum of 25% of a single image with all three labels to start, with further annotations on additional images added as necessary based on pipeline performance and to account for experimental group variability.

iii. Set location for exporting results under Prediction Export as outlined above.

b. If previous training needs to be changed, further annotations can be added or removed under Training in any of the three label classes.

Note: If the results are not satisfactory, review the uncertainty filter (chosen from the list under Training) and make annotations in regions with high uncertainty.

2. Import the experimental data (including any images for your dataset used in training) into Batch Processing under Select Raw Data Files and select Process all files, allowing the program to run.
3. Three versions of each input image will be exported to the location chosen in step 1. Probability Map Zero (Figure 2B) is what will be used for the following steps (named *title of image*_Probabilities_0).
   a. Review the quality of cell shape recognition of these probability maps by visually comparing the original input image and the probability map. Discard any images with obvious errors. See Troubleshooting problem 2.
   b. Ensure that the cells of interest have been accurately represented during segmentation.

Note: Common errors include light inconsistencies, obvious cell boundary thickening, and plate dots or scratches (examples in Figure 3).

Composite image creation in ImageJ/Fiji

Timing: 1 h

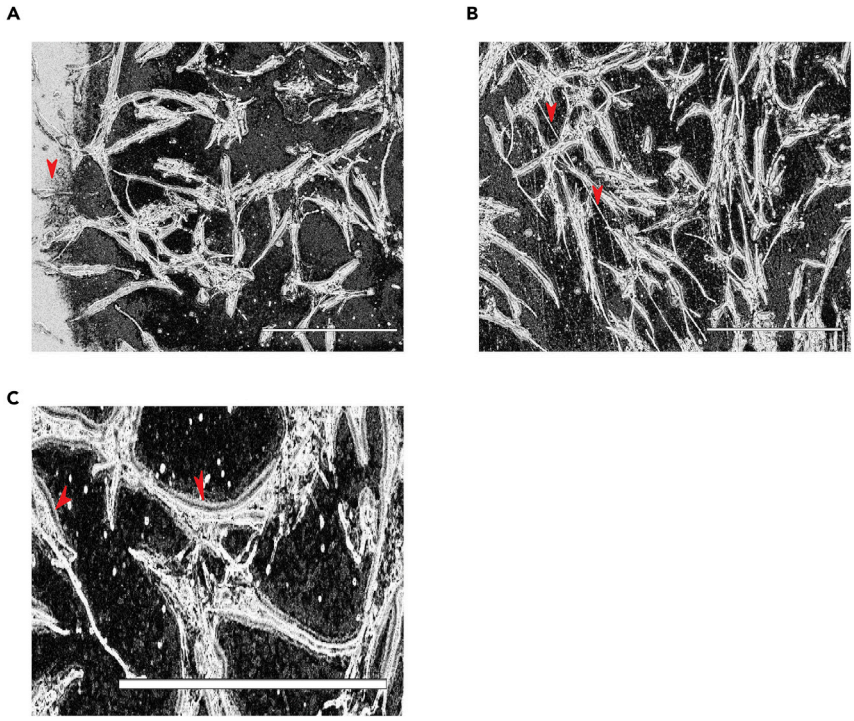


Figure 3. Common problems encountered during analysis (highlighted with red arrows)
(A) ilastik probability map with a lighting issue, causing a region to be improperly identified as part of the cell body.
(B) ilastik probability map with large scratch that would be misread as cellular projections.
(C) Individual cells from ilastik probability map with cell shape misrepresented, causing cells to appear thicker than actual size. Each scalebar is equivalent to 400 μm.

This section creates the images which will actually be quantified during analysis (Figure 2C). These composite images show both the cell body and nucleus, allowing for the nucleus to be used as a seed object by CellProfiler and ensure that every cell identified for measurement has an associated nucleus. This prevents the pipeline from arbitrarily drawing lines in between cells and eliminates the need for typical area to be used in cell identification.

4. Import the DAPI image and probability map of the first field into ImageJ/Fiji. Convert both images into 8-bit format (Image > Type > 8-bit).
   a. Threshold the DAPI image (Image > Adjust > Threshold) on the B&W default setting so that all nuclei are in black and the background is in white.

   Note: We recommend doing this by hand using the sliders in the pop-up window that ImageJ launches. Ensure that the dark background box is checked, and that the black background setting for binary functions is turned off (Process > Binary > Options > Black Background). If this setting is turned on, then the image will be inverted (with nuclei colored as white and background as black). The following steps will still function appropriately even if this inverted image form is used. See Troubleshooting problem 4.

   b. Use the Watershed function to separate overlapping nuclei (Process > Binary > Watershed).
5. Overlay the binary nuclei image with the probability map (Image > Color > Merge Channels). Set the DAPI image as the gray channel and the probability map as blue.
6. Convert the resulting image from 8-bit to RGB (Image > Type > RGB Color) to 8-bit (Image > Type > 8-bit), making it black and white. Save this image for analysis.

Note: The image should have the cell bodies highlighted in gray, nuclei in white, and background in black (Figure 2C).

Morphological measurement in CellProfiler

Timing: 2 h

This section performs the quantification of cell morphology (Figures 2D and 2E) and exports the results as a spreadsheet. Many of these steps require visual inspection of the image quality and use of the Test Mode function. Much of the description in the following steps details the settings used within the modules in CellProfiler.

7. Open the CellProfiler program and open the project (File > Open Project).

Note: Ensure that image type is set to Grayscale image and that the intensity range is from image metadata and that images are not processed as 3D under NamesAndTypes. Under Metadata, ensure that metadata is extracted from file/folder names, with the source being the file name from all images and the data type being text.

a. The modules and settings used in this pipeline are outlined below, though more are available in the software.
   i. Smooth: This module blurs small objects in the image (such as cellular debris, which can sometimes be incorrectly recognized as part of the cell body) and fills small holes in the cell body. This module is necessary as it allows the skeletonization modules to function more effectively later on.

Note: In this module, the input image is selected as DNA, the smoothing technique is set to Median Filter, artifact diameter is not calculated automatically, and the typical artifact diameter is set to 5 pixels, which we found to be effective in producing connected skeleton branches in later steps.

   ii. IdentifyPrimaryObjects: This module identifies the nuclei highlighted in white in the composite image, so that they can later be used as seed objects. It is important to note that at this stage the program is only able to "see" the white nuclei due to the threshold settings. The input image is set to the previously smoothed image from step 7ai and the reference range of object diameter for detection is 3–40 pixels for IMR-90 cells, though this range may need to be adjusted depending on cell type and microscope used.

Note: While this pipeline is set to not discard objects even if they fall outside of the set range, the range is set broadly as a reference for the software to detect nuclei and can be narrowed as necessary (see Troubleshooting problem 4). Nuclei are also not discarded if touching the border of the image. The threshold strategy is Global Otsu three-class thresholding as used previously,[8] middle intensity pixels are added to the background, threshold smoothing is set to zero and correction to 1.0. The lower and upper bounds on the threshold are 0.8 and 1.0. No log transformation is performed before thresholding, and Intensity[8] and Propagate are used to distinguish clumped objects. All other settings except displaying local maxima are set to Yes. Holes are filled in identified objects after declumping only, and, if an excessive number of objects is identified, the pipeline is set to continue.

   iii. IdentifySecondaryObjects: This module identifies the cell bodies and associates them with the nuclei identified by the previous module (Figure 2D). At this stage, the program primarily identifies the gray cell bodies and black background.

Note: The input image is the original smoothed image from step 7ai, and the input objects are the nuclei from step 7aii. The secondary objects are identified by Propagation, and thresholding is done using Global Otsu two-class thresholding. Smoothing scale and correction factor are the same as the previous module, and the lower and upper bounds on the threshold are 0 and 0.1. There is no log transformation before thresholding, regularization factor is set to 0. All other settings are set to Yes.

iv. ConvertObjectstoImage: This module converts the cell bodies identified in the previous module to an image so that subsequent modules can process it.

Note: The input objects are the cell bodies from step 7aiii, and the color format is set to Binary.

v. MorphologicalSkeleton: This module skeletonizes the cell bodies, converting them into one-pixel branches (Figure 2E). This allows for branching measurements to be taken by subsequent modules. The skeletonization procedure (this and following modules) is adapted from prior work.[9]

Note: The input image is the result of step 7aiv.

vi. MeasureImageSkeleton: This module measures several branching pseudopodia characteristics for the total image.

Note: The input image is the result of step 7av.

vii. MeasureObjectSkeleton: This module quantifies branching pseudopodia per cell, including the number of endpoints (conceptualized as the number of pseudopodia) and the total length of all pseudopodia per cell.

Note: The seed objects are the nuclei resulting from step 7aiii, and the skeletonized image is the result of step 7av. The branchpoint image is not retained, and small holes are filled with a maximum size of 10 pixels. Filling is necessary because, according to the software developers,[4] the skeletonization process involves steps that can leave hole artifacts in the image, causing false branchpoints and trunks. We found a setting of ten pixels to best resolve this issue for our data set. Skeleton graph relationships are not exported.

viii. MeasureObjectSizeShape: This module performs morphological measurements, including area, perimeter, and FormFactor (circularity).

Note: The object sets chosen to measure are the cell bodies and filtered nuclei (as some were excluded due their cell bodies touching the edges of the image) from step 7aiii. Additional settings should dictate that Zernike features are not calculated, but advanced features are. Advanced features include statistics for object moments and inertia tensors calculated by CellProfiler for export in the resulting Excel sheets, and can be used for additional comparisons of cell characteristics between groups.[4]

ix. ExportToSpreadsheet: This module exports the results to an excel spreadsheet in a location chosen by the user.

b. Drag and drop all files for one experimental group into the images section.

Note: Add in and analyze only one experimental group at a time, as results will be exported together in a single Excel file.

8. Check the quality of the image analysis.
   a. Click the green Start Test Mode button at the bottom of the screen.
   b. Step through the modules using the Step button for each image. Move to the next image by clicking the Next Image Set button.
      i. The most important module to check is the IdentifySecondaryObjects module. Compare the colorful identification of the cells to the original brightfield image and ensure that the shapes of cells seen in the corresponding brightfield image (Figure 2A) are accurately depicted in the identification image (Figure 2D). Note which images are of acceptable quality. After ensuring that the results of this module are satisfactory, also check the IdentifyPrimaryObjects and MorphologicalSkeleton modules. Common issues can be found in the troubleshooting section.
9. Exit test mode and remove any images found to not be satisfactory for analysis.
10. Edit the ExporttoSpreadsheet module to reflect the desired save location.
11. Quantify images using the Analyze Images button and then repeat steps 7b–10 for any remaining experimental groups.

Note: It is helpful to keep record of which images are accepted and which are rejected at various stages of the pipeline and why. We found useful to document this using an excel files with columns for each stage.

EXPECTED OUTCOMES

The CellProfiler program yields a wide variety of morphological measurements and shape characteristics.[4] In our recent work, we have focused on cell surface area, perimeter, circularity (FormFactor), pseudopodia (endpoint) number, and total pseudopodia length.[1] Examples of the processing steps from which these outputs are measured are seen in Figure 2. Cell morphology characteristics have been shown to change under various conditions, such as chronic stress.[1] Protocols similar to this have also been used in other biological contexts and cell types, demonstrating the broad applicability of quantifying cell morphology.[8,9] Cell morphology phenotypes can be indicative of organism characteristics; for example, differences in cell morphology have been associated with donor age more closely than other established biomolecular measurements such as DNA damage response or adenosine triphosphate content.[10]

QUANTIFICATION AND STATISTICAL ANALYSIS

As mentioned in the above protocol, images of undesirable quality should be removed during the analysis, leaving only representative images with clear cells. After the data has been exported to a spreadsheet by the program, cells with overlapping nuclei must be removed, as they would be counted as single cells with twice the area. To do this and additional analysis, we perform the following steps. An outline of this section can be seen in Figure 4. Additional measurements can be extracted from the spreadsheets depending on interest.

1. Create an Excel file for the data. An example can be seen in Table 1.
2. From the spreadsheet containing nuclear aspects (named in step 7aiii, under "name the new primary objects" in CellProfiler), extract the nuclear FormFactor, Endpoint Number, and Total Skeleton Length.

⚠ CRITICAL: Ensure that the spreadsheet that you are accessing here is the one named in step 7aiii, which should include only nuclei whose associated cell bodies are NOT touching the edges of the image. There is another spreadsheet containing all identified nuclei that should not be used for analysis.

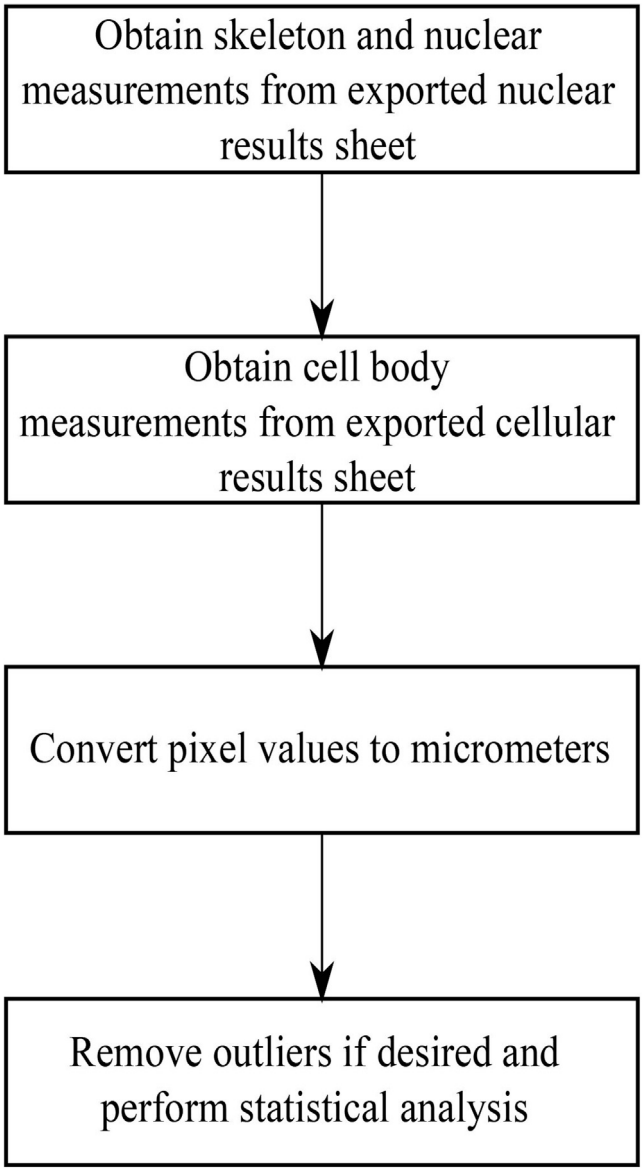


Figure 4. Simple schematic of quantification and statistical analysis

a. Paste these numbers into the spreadsheet in the corresponding columns, ensuring that they are in the exact order in which they appear on the spreadsheet. The total skeleton length should be added to the "Totalskeleton pix" column, as it will be converted later.

3. From the spreadsheet containing cellular aspects (named in step 7aiii), extract the Area, Perimeter, and FormFactor and place them into the spreadsheet.

Note: Make sure that you have separate columns for the cellular FormFactor and nuclear FormFactor.

a. Area and Perimeter should be pasted into the corresponding pix column. Ensure once again that all values are pasted in the same order that they occur in the original spreadsheets.

Table 1. Example of data formatting sheet

Treatment	Area (pix)	Area (μm)	Perimeter (μm)	Perimeter (pix)	Nuclear FormFactor	Cellular FormFactor	TotaleSkeleton (pix)	TotalSkeleton (μm)	Endpoints	LengthperEnd (μm)
Cortisol	2632	2264.09	314.672	365.806	0.848	0.247	181.238	155.903	8	19.488

4. Convert all pixel column values to micrometers. This conversion can be found using the line measurement tool in ImageJ/Fiji on the scale bar of one example image, which will give the length in pixels of the scale bar.
5. Divide the converted (micrometer) Total Skeleton Length value by the number of Endpoints for each cell, giving the average length per pseudopodium (endpoint) per cell.

   Optional: Duplicate the Excel sheet you are working on.

   Optional: In the duplicated Excel sheet, remove overlapping cell outliers.

   a. Operating under the assumption that the majority of cells are not overlapping within the dataset, use the nuclear FormFactor value as a metric of nuclei overlap.
   b. Calculate the inter-quartile range (IQR) of the nuclear formfactor values.
   c. Sort the data sheet according to nuclear formfactor (Sort&Filter > Custom Sort > Sort by Nuclear Formfactor). Delete rows with formfactor values outside of the IQR. Ensure that the entire row is deleted, effectively removing the cell from analysis.
6. Perform statistical analysis on the data as desired.

LIMITATIONS

There are several possible limitations to this protocol. Firstly, the current iteration of the protocol may have difficulty distinguishing very thin, nearly transparent cells from the background, as in such cases there may be little color difference to detect. Often, this results in the program missing large portions of thinly spread cells at the ilastik processing stage. This is something to control for when examining image quality at all stages, and the user should consider removing images in which this is a major issue. This issue may affect some cell types more than others. Secondly, while in our experience the pipeline consistently identifies cell pseudopodia, the user has little control over how the CellProfiler program quantifies individual pseudopodia. Thirdly, the pipeline cannot distinguish overlapping cell regions, as it utilizes two-dimensional images to represent three-dimensional objects. Therefore, the seeding density and confluency must be kept consistent during cell culture, so that the likelihood of cell overlap is minimized and comparable across experimental conditions. Finally, the circularity metric (FormFactor) is calculated using a ratio between the perimeter and area,[11] and thus cannot distinguish elongated from irregularly shaped cells. More advanced shape outputs are available from the CellProfiler program if necessary.[4]

TROUBLESHOOTING

Problem 1

When fixing cells, one experimental group appears to have washed away more easily than others. This can lead to both a much smaller pool of cells to image, and cellular debris that can be recognized as cell bodies by the pipeline, thus skewing analysis. Before you begin step 3f.

Potential solution

Decrease the amount of time that you place the cells in the freezer when fixing with methanol. This can make the cells less fragile during washes.

Problem 2

When processing images through ilastik, you see that several probability maps have large white areas instead of clearly defined cell bodies. Before you begin step 6, and main step 3.

Potential solution

Retrain the ilastik pipeline using all images. If the issue persists, consider removing images with lighting issues from the data set. The best way to prevent this from happening is to use a consistent light percentage on your microscope when taking the brightfield images.

Problem 3

The trained ilastik model does not seem to be recognizing cell shapes as precisely as you would like, and you have already tried to retrain the algorithm. Main step 1ai.

Potential solution

Change the sigma values in the ilastik program. Particularly, if you feel like the boundaries of the cell are too thick, increasing the edge value may be helpful. Try adjusting these values to see what is best for your data set. ilastik also suggests features based on your data set, which can be a good starting point.

Problem 4

Some DAPI staining spots are present outside of the nucleus, leading the program to inadvertently identify them as nuclei. This can result in the program associating cell bodies with these small spots, leading it to 'create' cells which are not really there. Main step 4a.

Potential solution

Go back to the ImageJ/Fiji portion of the protocol, and ensure that you are removing as many of these false nuclei as possible during thresholding. If the issue remains, the CellProfiler parameters can be edited. It is helpful to turn on the exclusion of nuclei outside of the given diameter range in the IdentifyPrimaryObjects module, and to raise this range enough to exclude these small dots. Ensure that you are not excluding any actual nuclei during this process.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the lead contact, Anthony Zannas (anthony_zannas@med.unc.edu).

Materials availability

This protocol did not generate new unique reagents.

Data and code availability

This protocol did not generate new data or code other than what is described in this manuscript.

ACKNOWLEDGMENTS

The following cell line was obtained from the NIGMS Human Genetic Cell Repository at the Coriell Institute for Medical Research: I90-83. This work was in part supported by a Summer Undergraduate Research Fellowship (SURF) to E.M.W.

AUTHOR CONTRIBUTIONS

Conceptualization, E.M.W., A.S.Z.; computational analysis and imaging, E.M.W.; cell culturing and treatment, O.K.; writing, review, and editing, E.M.W., O.K., A.S.Z.; funding acquisition and supervision, A.S.Z.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Leung, C.S., Kosyk, O., Welter, E.M., Dietrich, N., Archer, T.K., and Zannas, A.S. (2022). Chronic stress-driven glucocorticoid receptor activation programs key cell phenotypes and functional epigenomic patterns in human fibroblasts. iScience *25*, 104960.

2. Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., et al. (2019). ilastik: interactive machine learning for (bio)image analysis. Nat. Methods *16*, 1226–1232.

3. Rasband, W.S. (2011). Imagej (US National Institutes of Health). http://imagej.nih.gov/ij/.

4. Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. *7*, R100.

5. Crabbe, L., Verdun, R.E., Haggblom, C.I., and Karlseder, J. (2004). Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. Science *306*, 1951–1953.

6. Borel, F., Lacroix, F.B., and Margolis, R.L. (2002). Prolonged arrest of mammalian cells at the G1/S boundary results in permanent S phase stasis. J. Cell Sci. *115* (Pt 14), 2829–2838.

7. Sommer C., Straehle C., Koethe U., Hamprecht F.A. Ilastik: interactive learning and segmentation toolkit. Paper Presented at: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE; 2011.

8. Ramilowski, J.A., Yip, C.W., Agrawal, S., Chang, J.C., Ciani, Y., Kulakovskiy, I.V., Mendez, M., Ooi, J.L.C., Ouyang, J.F., Parkinson, N., et al. (2020). Functional annotation of human long noncoding RNAs via molecular phenotyping. Genome Res. *30*, 1060–1072.

9. French, S.L., Vijey, P., Karhohs, K.W., Wilkie, A.R., Horin, L.J., Ray, A., Posorske, B., Carpenter, A.E., Machlus, K.R., and Italiano, J.E., Jr. (2020). High-content, label-free analysis of proplatelet production from megakaryocytes. J. Thromb. Haemost. *18*, 2701–2711.

10. Phillip, J.M., Wu, P.H., Gilkes, D.M., Williams, W., McGovern, S., Daya, J., Chen, J., Aifuwa, I., Lee, J.S.H., Fan, R., et al. (2017). Biophysical and biomolecular determination of cellular age in humans. Nat. Biomed. Eng. *1*, 0093.

11. Goliwas, K.F., Richter, J.R., Pruitt, H.C., Araysi, L.M., Anderson, N.R., Samant, R.S., Lobo-Ruppert, S.M., Berry, J.L., and Frost, A.R. (2017). Methods to evaluate cell Growth, viability, and response to treatment in a tissue engineered breast cancer model. Sci. Rep. *7*, 14167.