

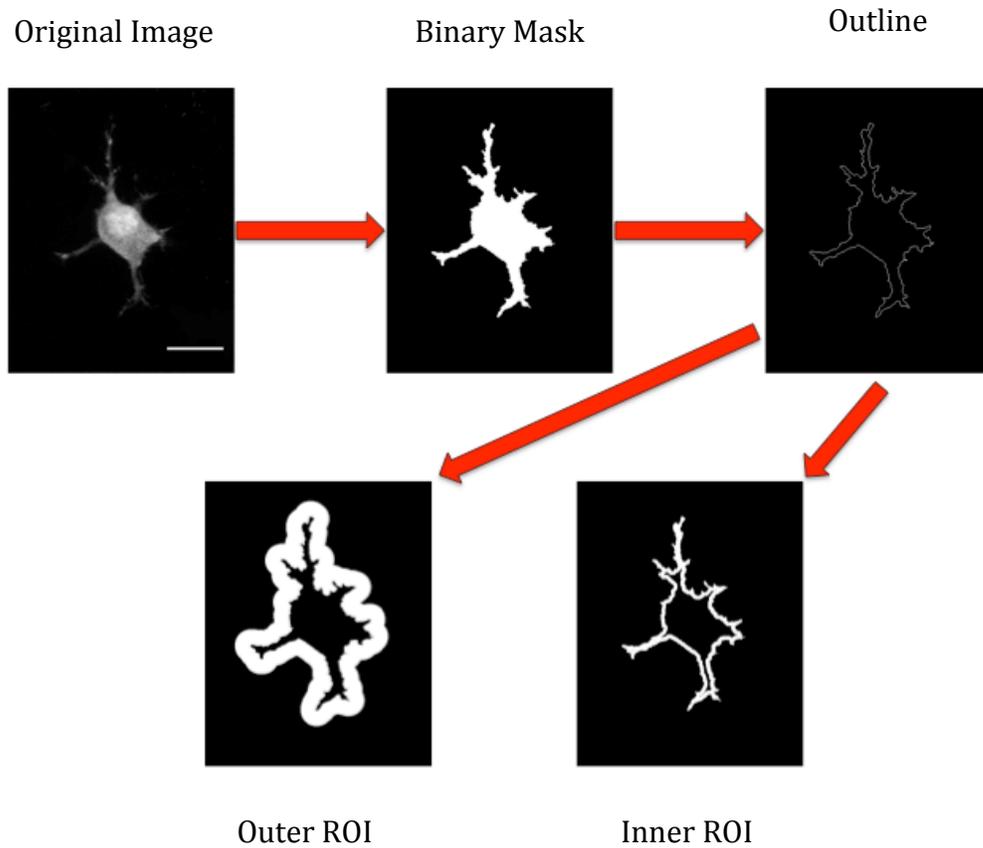
Description

The TumorTrace program is an automated image analysis tool, developed in MATLAB (The Mathworks, Inc., Natick, MA), for examining the ECM surrounding cells or tumors in the context of cellular morphology, protein expression and movement. It takes as input multiple image channels, either single images or stacks representing time-series or 3D data. It then finds a metric for the cell/cell cluster morphology and outputs plots representing intensity, morphology, collagen fiber alignment, and cell movement; .csv files containing raw data and image files containing the regions of interest.

Up to 5 images or image stacks may be uploaded via the GUI, at which point the user is able to assign a particular measurement to that image/stack (Morphology, Inner Intensity, Outer Intensity, Outline Intensity, or Alignment). The first measurement selected must always be Morphology, but the remaining channels may be assigned any measurement. The image channel that has been assigned the Morphology measurement is then segmented automatically using a histogram-based thresholding method. This segmentation is then used to create a binary mask from which a single-pixel width outline of the cell or cell cluster is generated. Then, for each pixel in this outline, the Euclidean distance between the outline pixel and the centroid of the cell or cell cluster is found. This is the surrogate metric used to describe the cell/cell cluster's morphology.

Three possible regions of interest may then be created, according to the user's measurement selection. The outline ROI and inner ROI are used for intensity measurements, while the outer ROI may be used for either intensity or fiber alignment measurement. Intensity at the outline ROI is calculated by applying the single-pixel outline to the desired image channel and averaging the 8-connect neighborhood surrounding each pixel in the single-pixel width outline. The inner ROI is designed to capture the intensity of labeled proteins just inside the boundary of the cell/cell cluster and its width is based on the size of cell/cell cluster. The intensity values for the inner ROI are calculated by applying the inner ROI to the desired image channel and finding

the average intensity inside the region using an averaging filter with kernel size equal to one half the width of the inner ROI. Intensity measurements in the outer ROI are performed in the same way, with the exception that the width of the outer ROI is a user-specified value (in pixels).



Operating Instructions

1) Launch program:



Channel 1

None

Name:

Load File

Channel 2

None

Name:

Load File

Channel 3

None

Name:

Load File

Channel 4

None

Name:

Load File

Channel 5

None

Name:

Load File

Compare Curves

Timeseries Analysis

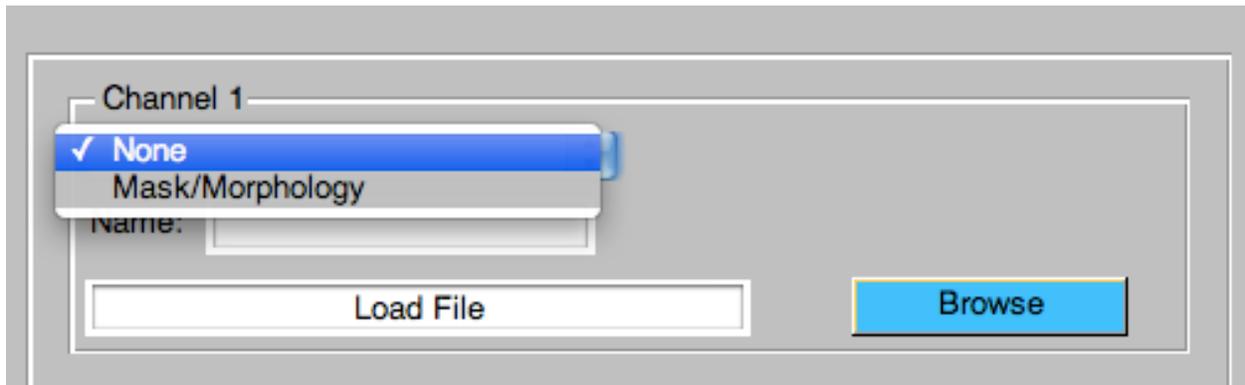
Enter ROI width in pixels:

Enter desired number of ROIs:

Intensity Threshold (0.05 default)

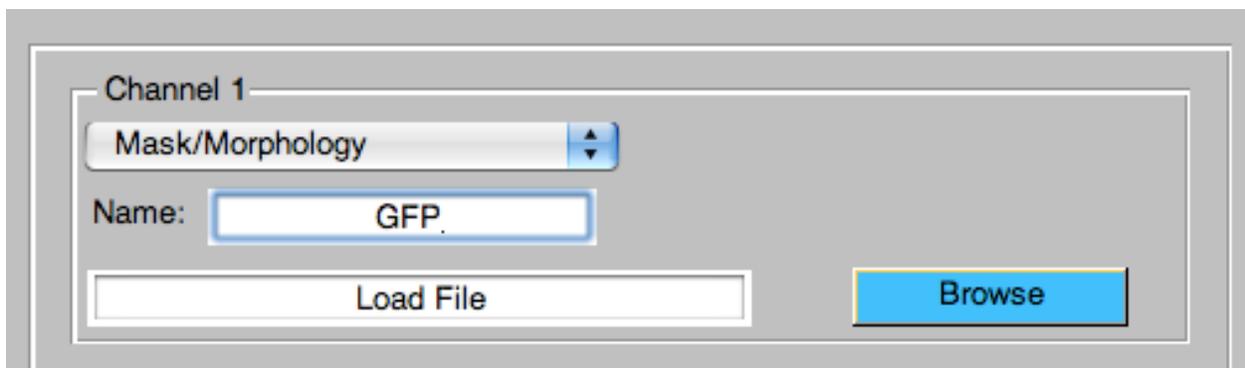
Curvelet Coefficient Threshold

2) Load images. This program allows the user to load up to 5 different images/stacks. To load an image/stack, first select the desired measurement for that image channel (the first channel must always be 'Mask/Morphology'):

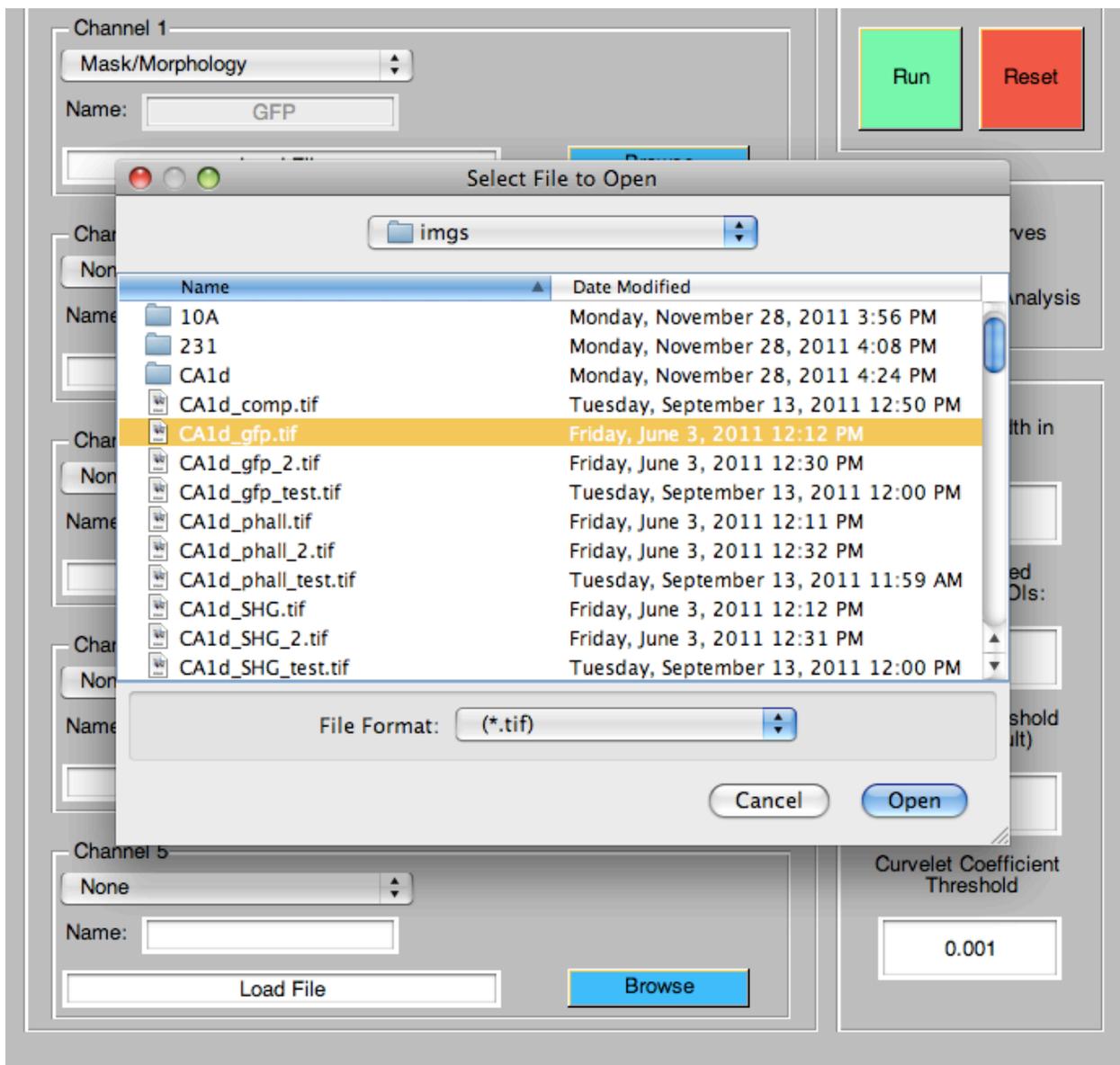


** The image used for 'Mask/Morphology' can be either a binary mask created in another program, such as Fiji, or an image of your cells with an evenly expressed protein marker or stain. **

Next, type in a name for this image channel. This name will be used in plot legends and filenames.

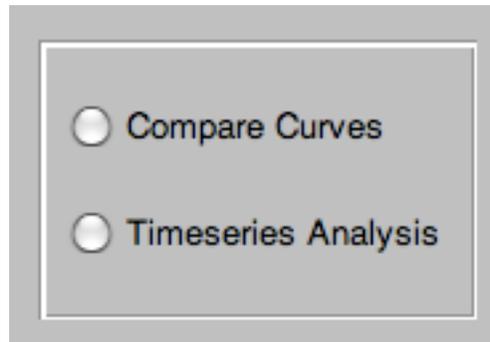


Finally, load the desired file using the 'Browse' button.



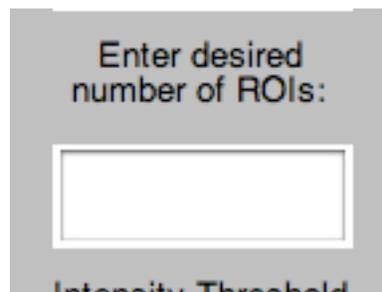
Repeat for remaining channels as desired.

3) Choose analysis type. The two options are: Compare Curves and Timeseries Analysis. Compare Curves performs analysis around the entire cell outline at each time point or z-plane, whereas Timeseries Analysis is the analysis of small regions of the cell outline through all time points or z-planes.



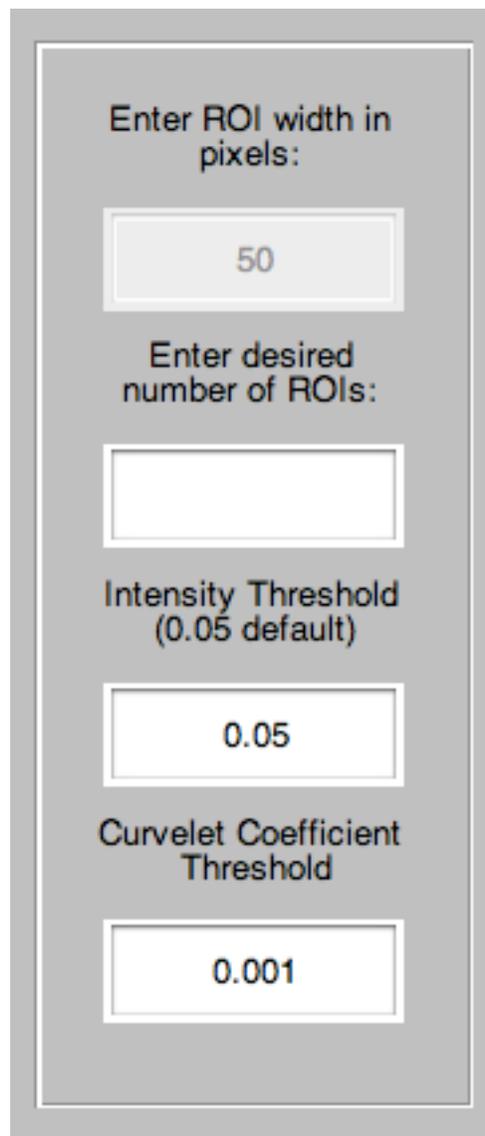
A gray rectangular panel containing two radio button options. The top option is "Compare Curves" and the bottom option is "Timeseries Analysis". Both radio buttons are currently unselected.

If Timeseries Analysis is chosen, the desired number of regions along the cell outline must be entered. A good rule of thumb for best performance is to try to keep the number of regions x the number of timepoints or slices under 200.



A gray rectangular panel with the text "Enter desired number of ROIs:" above a white rectangular input field. Below the input field, the text "Intensity Threshold" is partially visible.

4) Enter the width (in pixels) of the outer ROI, the intensity threshold, and the curvelet coefficient threshold. The intensity threshold is used in the creation of the mask. The default value, 0.05, works well for cytoplasmic GFP and other high contrast markers. The curvelet coefficient threshold is used only when collagen alignment in the outer ROI is performed. The larger the value, the more fibers will be measured.



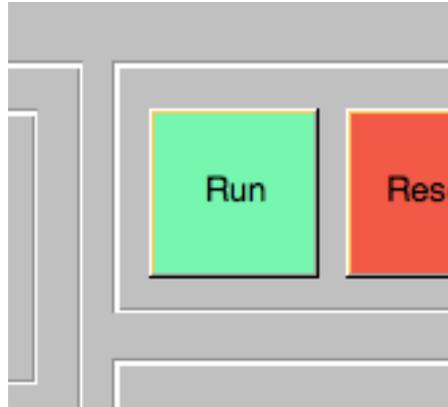
Enter ROI width in pixels:

Enter desired number of ROIs:

Intensity Threshold (0.05 default)

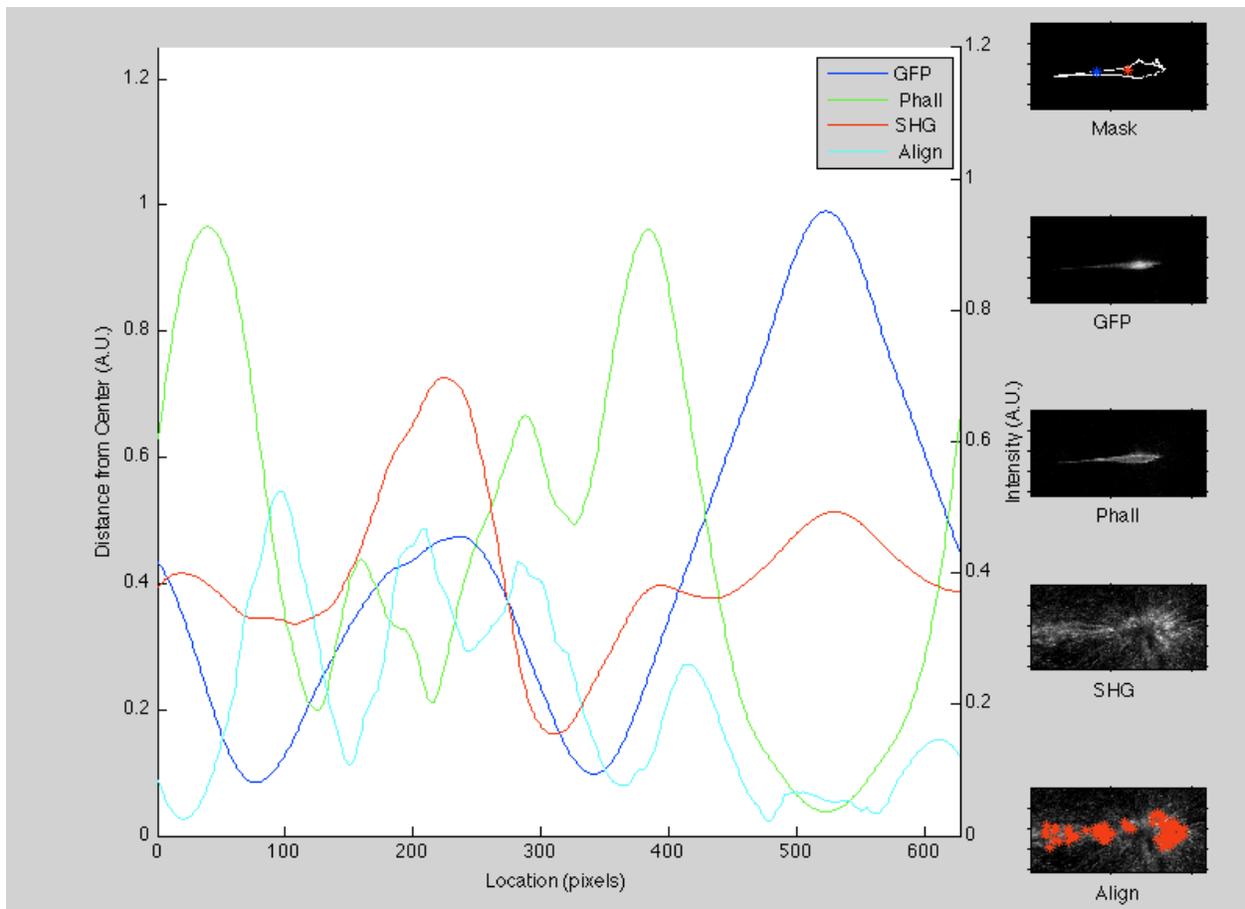
Curvelet Coefficient Threshold

5) Run program. When the 'Run' button is clicked, you will be prompted to select a directory for the output.



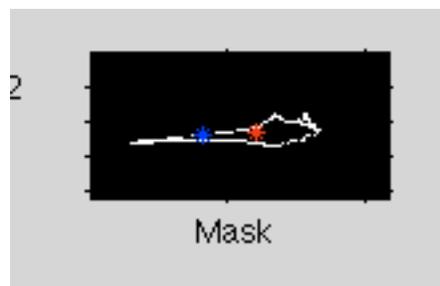
6) Results. The TumorTrace program provides the following results:

Plot of the measured intensities, morphology, and/or alignment at each time point or z-plane:

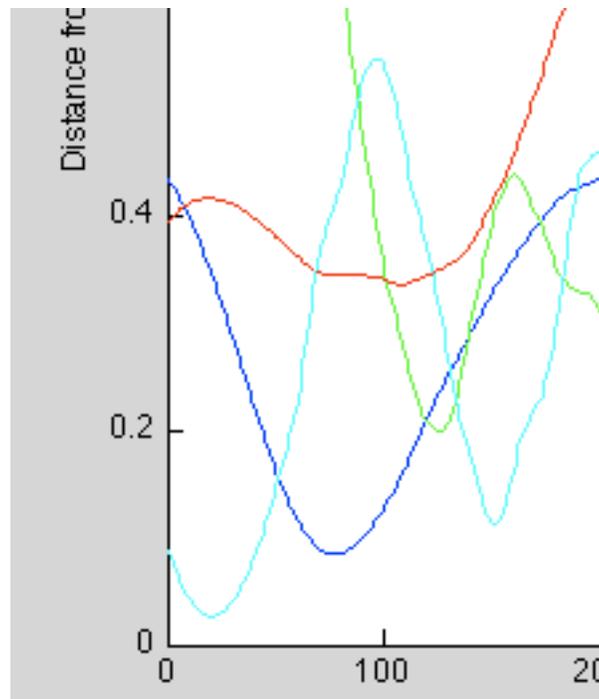


Interpreting the plot:

Each pixel of the single-pixel-width outline is represented as a location (x-axis). The starting point of the plot is the blue star on the Mask image in the upper right corner:



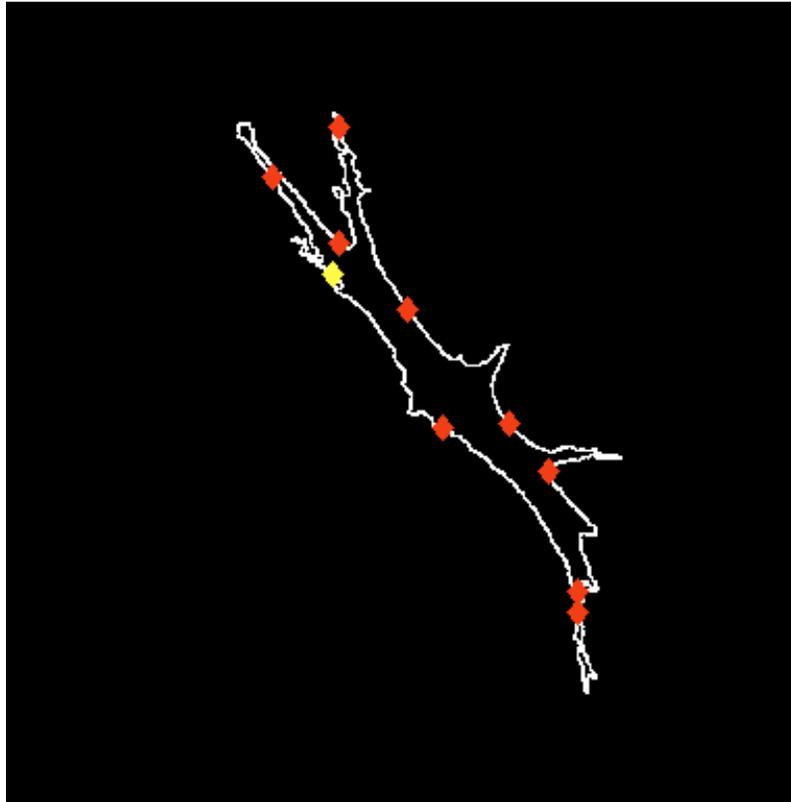
This point corresponds to the first pixel location on the plot:



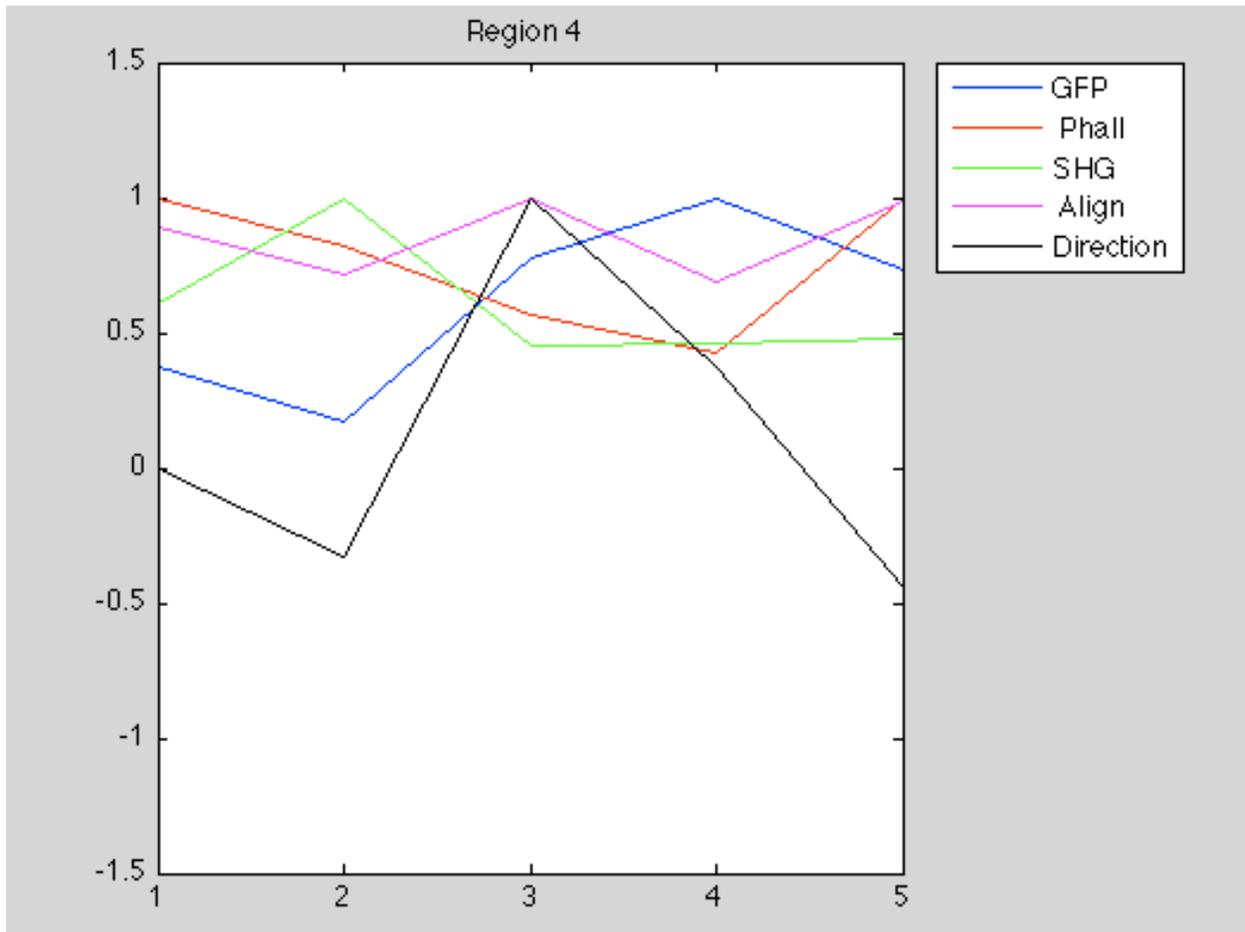
The plot then follows a path clockwise around the cell, giving the distance from the cell center (Morphology), intensity of labeled proteins or SHG, or collagen fiber alignment corresponding to each pixel.

The alignment value on the plot is normalized to 90° , so a value of 1 on the plot means that the fibers around that point are exactly perpendicular to the cell outline.

Plots of measured intensity, morphology, and/or alignment for each region of the cell outline through all timepoints (only if Timeseries Analysis is selected).



This example cell outline has been divided into 10 regions for Timeseries Analysis. The yellow diamond indicates the starting point, and the regions are delineated by red diamonds. The regions are numbered going clockwise around the cell, starting at the yellow diamond.



A plot of the measured values from each region of the cell outline through all timepoints is generated (x-axis is time, y-axis is normalized value). The additional metric, Direction, is added for timeseries data. A positive value indicates that the region is farther away from the cell center than it was in the previous timepoint, while a negative value means that the region is closer to the cell center than it was in the previous timepoint.

.csv files of all raw measurement data.

The data from each measurement channel will be saved in a .csv file with the name that you selected when loading images. The data in these files are organized so that each **row** represents one image (time point, z-slice) and the **columns** are the values.

	A	B	C	D
1	127.55	126.8	126.05	125.0
2	131.66	130.9	129.92	128.7
3	65.85	65.037	64.054	62.89
4				
5				
6				
7				

If Timeseries Analysis is selected, the .csv files containing the timeseries data will be named 'timeseries_region_x'. Each **row** represents the measurement channel, in the order that they were selected. Each **column** contains the value of the region at a given time point.

	A	B	C	D	E
1	83.963	38.352	175.33	225.89	165.02
2	27.919	23.002	15.778	11.923	27.775
3	37.313	61.169	27.687	28.027	29.033
4	59.902	48.063	67.148	46.242	65.978
5	0	-45.611	136.98	50.555	-60.868
6					
7					
8					

.csv files containing Spearman Rank Correlation Coefficients

The Spearman Rank Correlation Coefficient is calculated between each variable, and the results are written to .csv files named 'SRCC_Name1_Name2'. Each value in the file represents the Spearman Rank Correlation Coefficient between variable 1 and variable 2, over the entire cell outline for each time point or z-plane.

	A	B
1	-0.15716	
2	0.034997	
3	-0.029808	
4	0.00069063	
5	0.02931	
6		
7		

If Timeseries Analysis is chosen, the Spearman Rank Correlation Coefficient between each variable is calculated for each region, through all time points and the results are written to a .csv file named 'SRCC_time_name1_time_name2'. Each value in the file represents the SRCC between variable 1 and variable 2, at each region of the cell outline.

	A	B
1	-0.4	
2	0	
3	-0.4	
4	-0.4	
5	-0.8	
6	-1	
7	-0.4	
8	-0.4	
9	-0.4	
10	0.2	
11		
12		