WHISIT
WHere-IS-IT

A simple program for quantitative analysis of the fluorescence signal in populations of cells

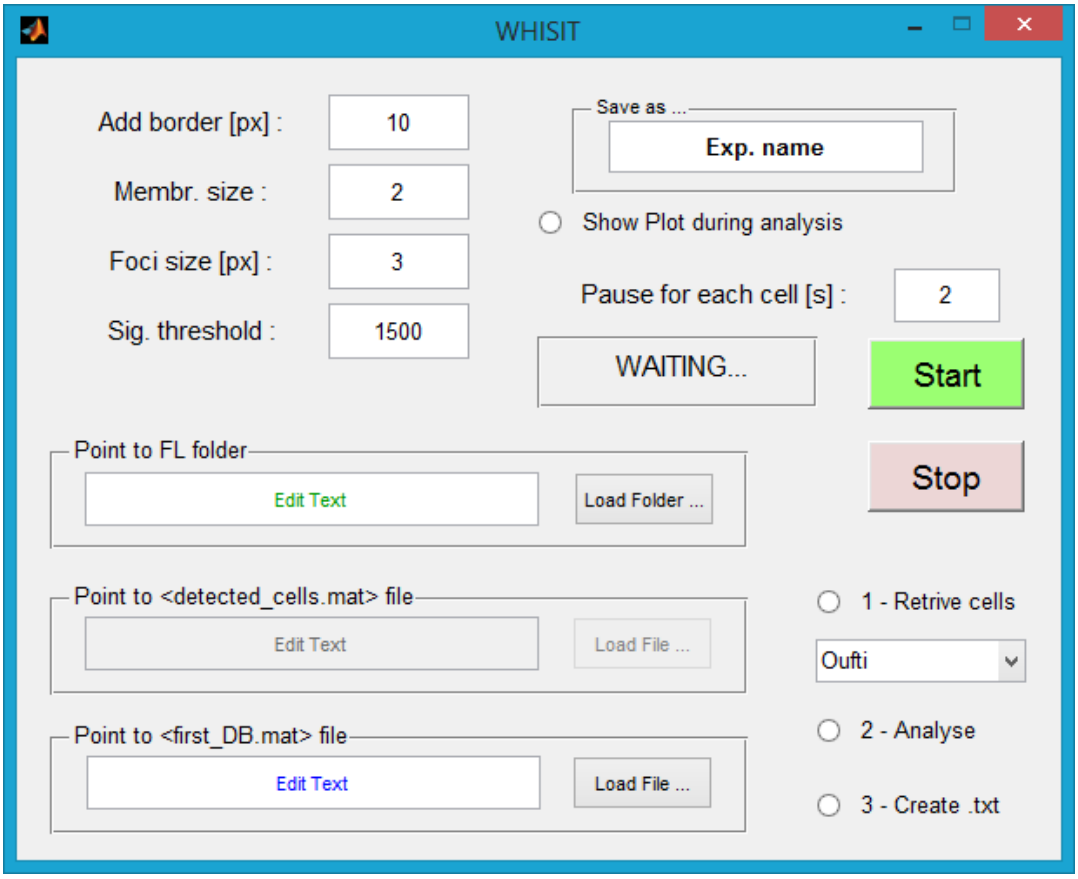
Table of Contents

Developed by Matteo Sangermani member of the Jenal Lab University of Basel, Switzerland (17/09/2016).

OVERVIEW

There are several programs for the analysis of foci and count number of spots in bacterial cells. But there is basically none that is addressing the issue of the ratio of signal between cytosol and membrane or showing the signal profile along a cell cross-section. For this reason the program WHISIT was created.

WHISIT uses fluorescence microscopy images and meshes of the cell outline computed for the corresponding bright field image by either MicrobeTracker (http://microbetracker.org/, (1)) or Oufti (http://oufti.org/index.html, (2)). It divides the cell into membrane and cytosolic compartment and performs quantitative analysis of the fluorescence signal for both compartments. Additionally, it computes the fluorescence intensity profile along cross sections of the cells. The cross sections are laid through foci detected in the cell.



N.B.: This program was tested with fluorescence microscopy images taken with a DeltaVision core microscope and SoftWoRx software which were further processes using ImageJ (FIJI, (3)) software and saved as .tiff files. Functionality cannot be guaranteed with images from different sources.

HOW TO USE IT

Input parameters

- ***Cell Box***
  When analyzing a cropped image with a specific cell in the center is considered. This value determines the extra border to add at the cropped picture of the cell. It does not change analysis results and can been increased if necessary, but should be at least >10 or there might be errors. (

  N.B.: the bigger they are the more memory the DB will use

  N.B.: if on the fluorescence image the cell is closer to the image border as the number of pixel added here, overlay of the computed cell outline and the fluorescence image does not work properly

- ***Membrane size***
  This value defines the thickness of the membrane area. For each increment the membrane is increased in width inward to the cytosol by an additional 1 pixel. Therefore, a value of 0 mean that the membrane width is just 1 pixel, corresponding to the outline of the cell perimeter calculated with MicrobeTracker/Oufti. A value of 2 results in a width of 3 pixels.

  N.B.: recommended values are around 1-3. Although physically the membrane is at most few tens of nanometers, the emission of fluorescent signals is still 400-700 nm. This should be considered to cover the membrane area correctly.

- ***Foci size***
  This value determines the minimal area (in pixels) that foci must have to be evaluated. Below such a value the detected foci are discarded as noise.

- ***Signal threshold***
  This is the threshold value that is applied to detect foci. Above such a value the pixels are examined for pattern and clusters and, if they are above the foci size, they are examined and evaluated.

  N.B.: Ii this value is too high in comparison to the signal intensity in the images, it can happen that WHISIT will save empty output files. In that case the signal threshold has to be decreased.

Final Remarks:

use the “Show Plot” during analysis to check if the signal is done properly. It is possible that a very a stringent cell detection in MicrobeTracker/Oufti, leaves very “thin” cells that have some of the fluorescent signal “out” of the cell and not considered for analysis. Perform again analysis in MicrobeTracker/Oufti, having more relaxed parameters.
A very good alignment of the fluorescent and bright field images is necessary. The program does not account and does not correct for possible shifts.

A 3-step process

1. ***Retrieve Cells*** – retrieves the cell detection analysis from the <detected_cells.mat> file and creates a new DataBase format (<Exp_name>*_1_DB.mat*) necessary for step 2.

    - *Popup menu*: choose whether the input <detected_cells.mat> was originated using Oufti or microbeTracker (although similar, the data is structured differently).
      (**N.B.**: That is why output *_1_DB.mat* is generated, because it has always the same format, independent of the program used for cell detection.)

2. ***Analyze*** – the main function that performs the analysis of the fluorescent signal for each cell. As inputs you need to provide the folder where fluorescent images are located and the *_1_DB.mat* file, which contain the data about the outline of each cell in the bright field.

    - (**N.B.**: output *_1_DB.mat* can be used over and over again and generate all possible results desired, with no change to it, since a new output *_2_DB.mat* is created at each analysis)

3. ***Create.txt*** – extracts the data from output *_2_DB.mat* and saves it as 2 .txt file that are ready to be used in Excel.

STOP button: works only during step '2 – Analyze'. This is the slowest step and STOP button allows to interrupt analysis at any point. However, data analyzed thus far is lost. If you push START, it will begin from the first cell in the list *_1_DB.mat*. Data is saved only at the end of the step '2 – Analyze'.

Step 2 can be sped up by either unchecking the field "show Plot during analysis" (fastest) or by decreasing the time a cropped image is shown during analysis by changing the parameter of "Pause for each cell [s]".

N.B.: Step 2 and 3 can be ticked together and WHISIT will perform both actions as soon as START is pressed.

N.B.: After data analysis and .txt file creation is completed in the WHISIT interface there are still some calculations ongoing in the background which can take some time, especially when a lot of cells or many foci are detected. If during this time a new analysis is started, WHISIT will not be able to save all data in the .txt files.

Folders and Images naming

The folder of each experiment/strain can be named as desired. However, it is very important to keep the naming of images consistent with the following rules.

The images must be have ***prefix***, "_" score symbol ad then followed by a 4 digits ***number***. The number must start at one and has to be continuous; fill with zero(s) if numbers have less than 4 digits (i.e. FL_0001, …, FL_0009, …, FL_0142, …). The fluorescent frames must be in the same order as bright field frames. Hence, it is recommended to name the bright field images in a similar way (e.g. BF_#### and FL_####). See sample images for an example

N.B.: ImageJ can easily change naming. Open the images as a single stack and then Save as → Image Sequence.

Why this weird arrangement? The additional zeros ensure that images in each folder are read in proper incremental order, which is recognized across all programs used (microbeTracker, Oufti and WHISIT). This in turn ensures correlation between the image numbers of your bright field and fluorescence images, regardeless of the program. If this is not done the program still works, but due to mismatches in bright field and fluorescence weird results are generated.

File Output

The outputs generated by the three steps are automatically named as follows. IMPORTANT: *all files are saved in the folder where the fluorescence picture folder is located*.

- ***<Exp_name>_1_DB.mat***: the new DataBase format
- ***<Exp_name>_2_DB.mat:*** the final DataBase with analysis results
- ***<Exp_name>_3_Cell.txt*** and/or ***<Exp_name>_3_Foci.txt***
  data extraction from <Exp_name>*_2_DB.mat* and rearranged for use in a spreadsheet.

In the second picture of Appendix there is an example of how the folders and file are recommended to be organized!

N.B.: file naming, such as *_1_DB.mat* or *_2_DB.mat*, can be changed after analysis. This do not compromising the possibility to use them it in WHISIT

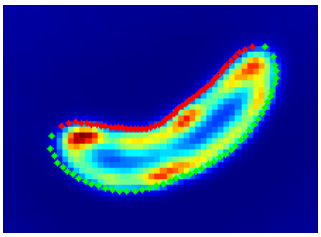
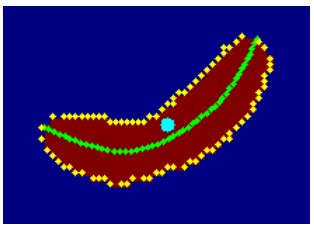
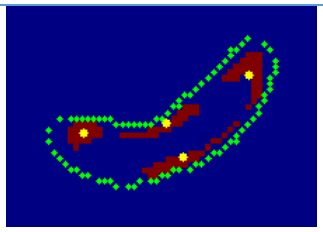
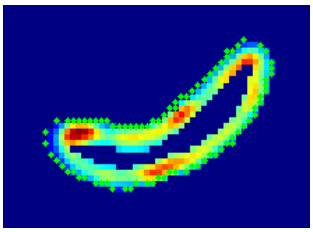
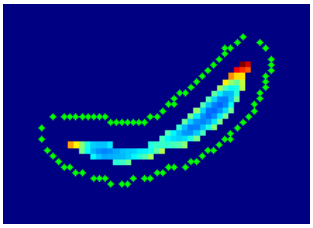
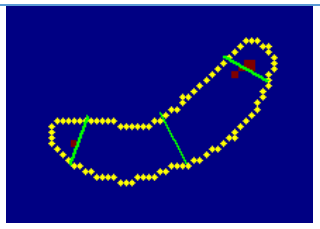
N.B.: if a value cannot be calculated (cell has no or too little fluorescence) some values are not calculated. Those will appear as **NaN**.

We named them signal clustering as "**foci**" for simplicity, but we do not use the word "foci" in the classical definition, as round spot of signals. Here with "foci" we are considering *a clusters of signal above a specific signal threshold level*. Shape is not important, so they can have any shape.

If we were to take the cell signal profile at predefined locations (i.e. cutting in the middle of the cell) this could be uninformative or just wrong. Therefore, we find all areas where the signal is clustered the most ("foci"), and have signal profile perpendicular to the cell axis and passing through their center.

THE ALGORITHM - HOW IT WORKS

The algorithm computes the fluorescence signal in the cytosol and membrane and the fluorescence signal profile along the cell width.

Description	Step
The cell outline identified in the bright field image is used as a mask to isolate the fluorescence signal of the specific cell. A cropped image of the cell of interest is generated and used for all subsequent calculations.	
The cell mesh is a set of coordinates that identifies the cell perimeter. This is then converted in pixel coordinates that define the cell perimeter. The cell longitudinal axis is also calculated from the mesh coordinates. Finally the area and center of mass of the cell are calculated.	
The next step involves identification of signal clusters (foci) using ***Signal threshold*** and ***Foci size*** as criteria. For each of those clusters a mask is generated and stored and used to calculate the average signal intensity, areas, etc…	
The algorithm then identifies the membrane and cytosol area according to the parameter used. For both a mask defining the area in the cells is generated and stored. The ***Membrane size*** is the main parameter that defines the final extent of each area.	 
For each foci a profile line is created. This line runs through the center of the foci and is orthogonal to the cell axis. The profile line is an array of pixel values in the place it crosses the cell	

OUTPUT FILE FORMAT

There are 2 ***.txt*** files generated with tab-separated columns.

_Cell.txt

Data is collected from the cell point of view: ***each row is a cell entry*** with all related information about a specific cell

	Column value	Description
1	***Cell Length***	Length of cell longitudinal axis [pixel]
2	***Number Foci***	Number of foci identified in the cell
3	***Abs tot Foci_A in Memb_A***	Fraction of total area of all foci that reside in the membrane area
4	***Abs tot Foci_A in Cyto_A***	Fraction of total area of all foci that reside in the cytosol area (column 3 + 4 must give a result of 1)
5	***Rel Foci_A in Membr_A***	Fraction of membrane area that is occupied by Foci
6	***Rel Foci_A in Cyto_A***	Fraction of cytosolic area that is occupied by Foci
7	***Avg Foci eccentricity***	Average eccentricity of foci. How round an ellipse is: value of 0 is a perfect circle, value of 1 an infinitely stretched ellipse.
8	***Avg Foci Area***	Average area of the foci present in the cell [pixel^2]
9	***Avg Foci distance***	Average distance between the foci center and cell center of mass [pixel]
10	***Area Membr***	Area of the membrane [pixel^2]
11	***Area Cyto***	Area of the cytosol [pixel^2]
12	***Area Cell***	Area of the whole cell (membr + cyto) [pixel^2]
13	***Avg Int Membr***	Average intensity value of the signal in the membrane area
14	***Avg Int Cyt***	Average intensity value of the signal in the cytosol area
15	***Avg Int Cell***	Average intensity value of the signal in the cell area
16	***Avg_BkGr_noise***	Average pixel value in the background area for the frame in which the cell reside
17	***Min_px_value_Frame***	Lowest pixel value in the whole frame in which the cell belongs to

N.B.: Basically both Abs and Rel value tell us where the foci (cluster of signal) preferentially reside in a cell; however, the Abs values tell us how foci are split between membrane and cytosol, while the Rel values tell us how much of each part (membr. and cyto) is occupied by foci.

Neither of ***Avg Int Cyt***, ***Avg Int Membr*** or ***Avg Int Cell*** use the “Foci” for calculating their value.

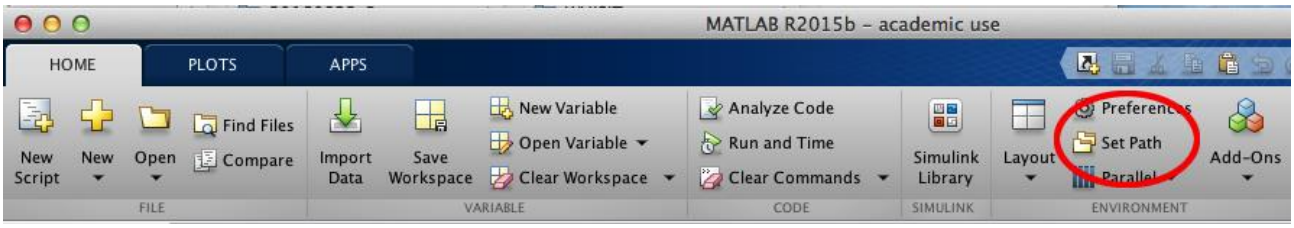
_Foci.txt

Data is collected from the Foci point of view: ***each row is a focus entry*** with of all information concerning that one focus. Many cells have multiple foci and each will be treated independently

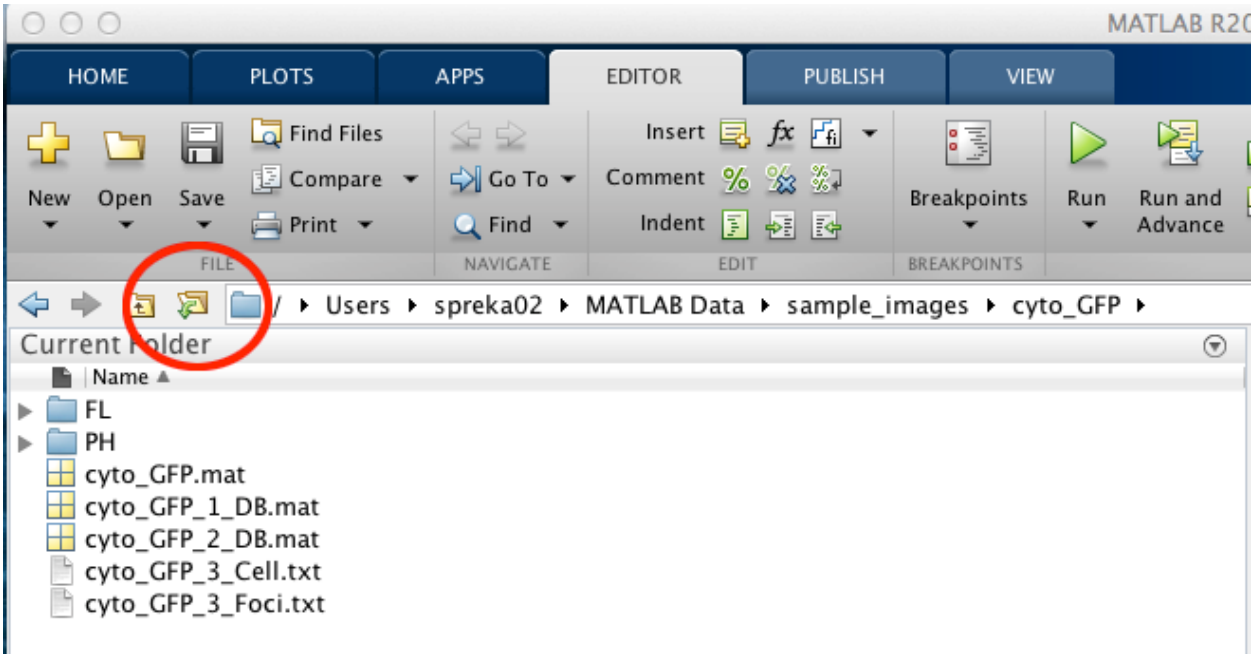
	Column value	Description
1	***Foci area***	Area of the focus
2	***Abs Foci_A in Membr_A***	Fraction of area of a specific focus that resides in the membrane area
3	***Abs Foci_A in Cyto_A***	Fraction of area of the specific focus that resides in the cytosolic area (column 2 + 3 must give a result of 100%)
4	***Avg Int signal***	Average intensity of the signal in the focus
5	***Eccentricity***	Average eccentricity of the focus
6	***Distance from cell Center***	Average distance of the focus from the cell center
7	***Length cell it belongs***	Length of the cell the foci belongs to
8	***Avg_BkGr_noise***	Average pixel value in the background area for the frame in which the cell belongs to
9	***Min_px_value_Frame***	Lowest pixel value in the whole frame in which the cell belongs to
10	***EMPTY***	
11	***From 11 onward***	Signal profile perpendicular to the cell axis crossing the focus center. Each profile is normalized to be of length 25 (which is also the highest length/width possible), with additional zeros left-right in order to center the array. Since center is where the cell axis passes, this allows some degree of comparison between different signal profiles at different locations of the cell.
…	“ “ “	

APPENDIX: using MATLAB

MATLAB needs to know the “paths” for the folder it has to use. Therefore go on “Set Path”, then “Add with Subfolder” to add all script and functions of WHISIT.



It is also recommended to “Point to a folder” where your experiment’s pictures are located: the pointed folder is the one displayed when you need to input a path to a file/folder in WHISIT. If you point to it, you avoid digging every time the whole path for adding a file/folder and you can do it quickly in a couple of clicks.



APPENDIX: Sample images

The folder **sample_images** contains two sets of fluorescence images: Cells expressing eGFP, which localizes to the cytosol (cyto_GFP) and cells stained with the membrane dye FM4-64 (mem_FM4-64). Both sets contain the _1_DB.mat file derived from Oufti and the corresponding _2_DB.mat file computed by WHISIT. Additionally, for each set the output files _3_Cell.txt and _3_Foci.txt are provided which were computed with the following parameters: Add border 8, Membr. Size 3, Foci size 2, and Sig. threshold 450.

REFERENCES

1. **Garner EC**. 2011. MicrobeTracker: quantitative image analysis designed for the smallest organisms. Mol Microbiol **80**:577–579.

2. **Paintdakhi A**, **Parry B**, **Campos M**, **Irnov I**, **Elf J**, **Surovtsev I**, **Jacobs-Wagner C**. 2015. Oufti: An integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. Mol Microbiol 1–11.

3. **Schindelin J**, **Arganda-Carreras I**, **Frise E**, **Kaynig V**, **Longair M**, **Pietzsch T**, **Preibisch S**, **Rueden C**, **Saalfeld S**, **Schmid B**, **Tinevez J-Y**, **White DJ**, **Hartenstein V**, **Eliceiri K**, **Tomancak P**, **Cardona A**. 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods **9**:676–682.