Github Readme AH230823

This Matlab code was used in [article ref].

# Purpose of the code

These Matlab scripts allow to correlate nuclear signals in the green (ch2, “TAG” or T-antigen viral protein of interest) and red (ch3, “iPOND protein of interest”) channels using the Pearson Correlation Coefficient over a region which excludes nucleoli, in a subset of cells which display significant Viral Replication Centers. These scripts are amenable to high throughput datasets (z-stacks), but as presented, the analysis is manual (one image at a time).

Using Excel and ImageJ histogram analysis, we found that it is possible to dynamically delineate:

1. Viral Replication Centers (VRCs) which are TAg-rich regions in nuclei. The mathematical expression used for VRCs is the region where the nuclear ch2 signal is greater than a nucleus-dependent threshold. Namely: VRC > VRC\_factor \* median\_TAG(nucleus), where VRC\_factor is =2 by default. Namely, regions where the ch2 signal is greater than twice the median nuclear ch2 signal designates the VRCs at all stages of early to advanced viral infection.
2. Nucleoli, which are TAg-deprived regions in nuclei. They are also generally deprived in other markers such as the iPOND candidates. The mathematical expression used for nucleoli ‘holes’ is the region where the nuclear ch2 signal is lower than a nucleus-dependent threshold. Namely: nucleoli < Nucleoli\_factor \* median\_TAG(nucleus), where Nucleoli\_factor is = 0.6 by default. Namely, regions where the ch2 signal is lower than 0.6 times the median ch2 nuclear signal designates nucleoli at all stages of early to advanced viral infection.

Nucleoli typically appear as marker-poor regions, whether TAg or other iPOND proteins of interest. Therefore, in order to avoid an artificially high correlation between two diffuse uncorrelated markers yet both excluded from nucleoli, we wanted to define a PCC\_area to evaluate our correlation between ch2 and ch3 signals. Namely, our aim1C script evaluates a Pearson Correlation Coefficient (PCC) using the Matlab corr2 function over a nuclear area excluding nucleoli.

The ability to automatically demarcate VRCs in each nucleus allows us to estimate the VRC area nuclear fraction. In-turn, this could be used to discriminate between early-stage infection with more punctate morphologies (1-10% nucleus area), medium stage with larger centers (10-17% nucleus area as VRC), and late stage with tracks (>17%). In this work, we limit our analysis to TAg-positive cells which display clear VRCs, with VRC\_area\_fraction greater than 1% (by setting VRC\_fraction\_thresh = 0.01).

For each nucleus detected in ch2, the aim1C script:

- outputs a figure which delineates VRCs, nucleoli and the PCC\_area (figure used to tune the parameters VRC\_factor and Nucleoli\_factor)

- estimates the VRC area fraction in the nucleus

- If the VRCs are large enough (here, if VRC\_area\_fraction > 0.01), calculates the PCC value over the PCC\_area (shown in the figure bottom panels). If the VRCs are too small or absent, the script outputs "NaN" (not a number). The NaN output makes it easy to extract ‘not NaN’ values for PCC statistics restricted to VRC-positive cells.

# Producing and naming the Images for analysis

Images were acquired using the Perkin Elemer Opera Phenix.

For each iPOND protein of interest, 6-12 nuclei featuring significant Viral Replication Centers (VRCs, which appear as TAG-dense regions) were selected by cropping in ImageJ (see folder ‘Subset AH230226 github INPUT’) using ROI manager Rectangle selection. Note that for ch2 and ch3, the images for a given nucleus have the same dimension and overlay perfectly. Manual selection of the z-plane with maximum intensity was performed.

Example of file name: “r03c10f16p03-ch3sk1fk1fl1-3.tif”

Row03, column10, Site16, plane03, channel 3 [sk1fk1fl1]

“-3.tif” means that this is the third nucleus selected using ImageJ from the raw Phenix image named r03c10f16p03-ch3sk1fk1fl1.tif.

The code provides a means to find the name of the ch3 image provided the ch2 image name, even if they are not in the same z-plane noted as ‘p’.

# Workflow overview

Three scripts are provided to tune parameters critical for robust analysis. These parameters, which depend on the eukaryotic cell morphology, cellular response and the properties of the imaging system, can be adapted to different experimental settings with similar goals.

- Nuclear segmentation in aim1A,

- ch2/ch3 contrast parameters in aim1B,

- parameters to identify VRCs and nucleoli in aim1C.

Users can follow 1A => 1B =>1C, as detailed below.

Alternatively, they can directly use aim1C to tune the parameters by uncommenting the line imshow(Ci) by modifying relevant parameters.

A white background with black text

Description automatically generated

# Github\_Coloc23\_aim1A

## Goal: nucleus segmentation

As an alternative to using the ch1 DAPI/Hoechst staining, we use ch2 to delineate the nucleus. The goal of this script is to specify paths and adjust the following three parameters highlighted in yellow below. If segmentation fails, this script can be used to troubleshoot and the parameters fspecial\_radius and is\_WOP can be tuned.

## Paths

A screenshot of a computer program

Description automatically generated

## Input filename

Choose a few representative cells in dataset

A close up of a text

Description automatically generated

## Basic Parameters to adjust

% NUCLEAR SEGMENTATION PARAMETERS TO ADJUST

LowAreaBound = 300; % area debris in px^2 to exclude small objects

HighAreaBound = 50000; % area debris in px^2 to exclude large objects

% large objects exclusion not actually needed since the cells were already chosen

% manually

fspecial\_radius=6; %ADJUST default=6 for C57 MEFs cells, =2 for WOP cells

is\_WOP=0; % if you have WOP cells use =1; for C57 use =0 (default)

% ADJUST manually thresh\_bs - see section below

% "Image Overlay TEST" visualize bask\_bs1

thresh\_bs=600 ; % use about 2xmax value of empty region (image J histogram)

## Example outputs

GOOD

LowAreaBound = 300; HighAreaBound = 50000; thresh\_bs=600 ;

A screenshot of a computer screen

Description automatically generated

NOT GOOD (dots in the upper left corner)

LowAreaBound = 0; HighAreaBound = 50000; thresh\_bs=600 ;

A screenshot of a computer screen

Description automatically generated

NOT GOOD (cells are picked up as debris, need to increase HighAreaBound parameter)

LowAreaBound = 300; HighAreaBound = 5000; thresh\_bs=600 ;

A screenshot of a computer screen

Description automatically generated

NOT GOOD (the background threshold is too low in this case)

LowAreaBound = 300; HighAreaBound = 50000; thresh\_bs=200 ;

A screenshot of a computer screen

Description automatically generated

# Github\_Coloc23\_aim1B

## Goal: tune image contrast parameters in ch2 and ch3 images

## Copy from aim1A

A screenshot of a computer code

Description automatically generated

## Input filename

Choose a few representative cells in dataset

A close-up of a computer code

Description automatically generated

## Adjust parameters

The goal of this script is to adjust FOUR ch2 and ch3 image parameters (background\_ch2, background\_ch3, cutoff, gamma) such that the features of interest appear with enough contrast.

The nuclear mask is the same in all the figures, and was determined in aim1A. The color helps recognize the various channels:

Left panel, blue nuclear mask, ch2 or ch1 if available – not used in the aim1D output.

Middle panel, green nuclear mask, ch2 – this will be used in the aim1D output.

Right panel, red nuclear mask, ch3 – this will be used in the aim1D output.

% IMAGE DISPLAY PARAMETERS TO ADJUST

background\_ch2=300; % visualize output for infected vs health cells

background\_ch3=300; % visualize output for infected vs health cells

MinAdj\_ch2=1\*background\_ch2/65536;

MinAdj\_ch3=1\*background\_ch3/65536;

cutoff = 0.4; % 0.5 works if signal is dim in the channel; 0.4 for iPOND

gamma = 0.4; % default is 0.1 for immunoplaque; 0.4 for iPOND

## Example outputs

Example 1:

background\_ch2 = 300, background\_ch3 =300, cutoff = 0.4, gamma = 0.4

A group of white circles with green dots

Description automatically generated

Example 2:

background\_ch2 = 300, background\_ch3 =300, cutoff = 0.4, gamma = 0.6

A group of white circles

Description automatically generated

Example 3:

background\_ch2 = 300, background\_ch3 =300, cutoff = 0.4, gamma = 0.8

A group of white circles

Description automatically generated

In the example 3, the correlation between ch2 and ch3 is most apparent.

# Github\_Coloc23\_aim1C

## Goals: delinate VRCs, nucleoli, estimate PCC

See section “purpose of the code”.

## Copy from aim 1B

A screenshot of a computer program

Description automatically generated

A white screen with black text

Description automatically generated

A white screen with black text

Description automatically generated

## Input filename

A computer screen shot of text

Description automatically generated

## Parameters to adjust

For each nucleus detected in ch2, the aim1C script:

- outputs a figure which delineates VRCs, nucleoli and the PCC\_area (used to tune parameters VRC\_factor and Nucleoli\_factor)

- estimates the VRC fraction

- If the VRCs are large enough (here, if VRC\_area\_fraction > 0.01), calculates the PCC value over the PCC\_area (shown in the bottom panels). If the VRCs are too small or not present, the script outputs "NaN" (not a number). The NaN output makes it easy to extract ‘not NaN’ values for PCC statistics restricted to VRC-positive cells as defined by the parameter VRC\_area\_fraction

A computer code with text

Description automatically generated with medium confidence

## Example outputs

VRC\_factor = 2; Nucleoli\_factor = 0.6

A comparison of a person's body

Description automatically generated with medium confidence

VRC\_factor = 3; Nucleoli\_factor = 0.6

A comparison of a person's face

Description automatically generated with medium confidence

VRC\_factor = 1; Nucleoli\_factor = 0.6

A black and blue globe with white text

Description automatically generated

VRC\_factor = 2; Nucleoli\_factor = 1

A comparison of a blue and black image

Description automatically generated with medium confidence

VRC\_factor = 2; Nucleoli\_factor = 0.2

A comparison of images of a human skull

Description automatically generated

# Miscellaneous

The CORR2 function used in aim1C implements the Pearson correlation to 2-D arrays (source: <https://www.mathworks.com/matlabcentral/answers/102954-what-definition-of-correlation-is-implemented-by-corr2-function-in-image-processing-toolbox-6-4-r20>).