

Introduction to data analysis in R

Quantifying protein co-localization in fluorescence images using the colocr package

Mahmoud Ahmed

August 2, 2022

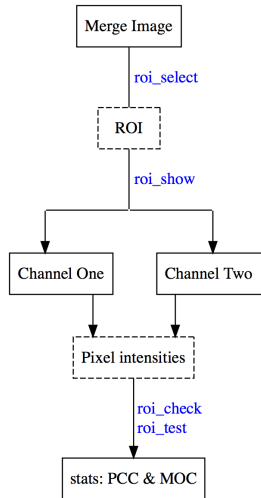


Fluorescence image analysis

A few R packages are available for conducting image analysis
The colocr package provides a simple straight forward workflow for

1. Loading images
2. Choosing regions of interest
3. Calculating co-localization statistics.

Included in the package, is a web app that can be invoked locally to interactively select the regions of interest in a semi-automatic way.

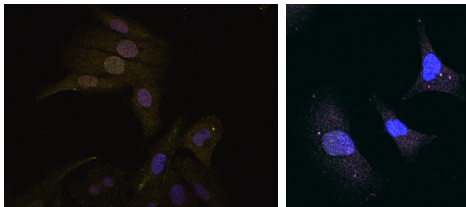


The dataset

These two images come from the DU145 prostate cancer cell line. Cells were stained with probes for two proteins RKIP and LC3.

The aim of this experiment is to determine, how much of the two proteins are co-localized or co-distributed in these cells.

The files containing these images are named 'image_1.png' and 'image_2.png'.



Loading images in R

We first need to load the images in R and look inside the object to see how R represents images in numbers and pixel coordinates.

```
# load colocr library  
library(colocr)  
  
# load image  
img <- image_load('image_1.png')  
  
# print img object  
img
```

Image. Width: 800 pix Height: 600 pix Depth: 1 Colour channels: 3

This is a four dimensions object of class cimg that contains the image data.

Image representation as numbers

To get a deeper understanding of this object, we can transform it into a data.frame and have a look at it.

```
# transform the object into a data.frame and call str  
str(as.data.frame(img))
```

```
'data.frame':  1440000 obs. of  4 variables:  
 $ x      : int  1 2 3 4 5 6 7 8 9 10 ...  
 $ y      : int  1 1 1 1 1 1 1 1 1 1 ...  
 $ cc     : int  1 1 1 1 1 1 1 1 1 1 ...  
 $ value: num  0.0392 0.0392 0.0392 0.0353 0.0353 ...
```

Here is a breakdown of the dimensions:

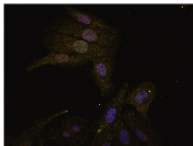
- x and y are the spatial coordinates of the pixels
- cc is the color (channel)
- value is the intensity

Visualizing images

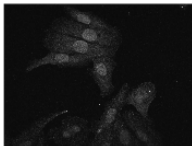
Channels 1, 2, and 3 refer to the red (RKIP), green (LC3), and blue colors (DAPI).

```
# show channels  
library(imager) # to all channel  
plot(img, axes = FALSE, main = 'Merged')  
plot(channel(img, 1), axes = FALSE, main = 'RKIP')  
plot(channel(img, 2), axes = FALSE, main = 'LC3')
```

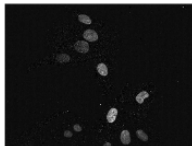
Merged



RKIP



LC3



Selecting regions of interest

To select a region of interest (ROI), we use `roi_select`. By default, the largest contiguous region of the image is selected.

Thresholding means all values below that number are set to 0, and above it to 1.

```
# select the region of interest  
img_rois <- roi_select(img, threshold = 90)  
str(img_rois)
```

```
'cimg' num [1:800, 1:600, 1, 1:3] 0.0392 0.0392 0.0392 0.0353 0.0353 ...  
- attr(*, "label")= num [1:480000] 0 0 0 0 0 0 0 0 0 0 ...
```

The returned `cimg` object contains the original input image and an added attribute called `label` which indicates the 0 (not-selected) or 1 (selected) for each pixel.

Visualizing selected regions

To make sure the selection is appropriately encompassing the ROI, call `roi_show`, which displays

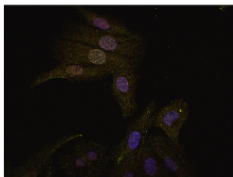
- The original merge image
- A low-resolution picture of the ROI
- Highlighted ROIs in different channels

```
# select ROI and show the results  
roi_show(img_rois, ind = c(1, 2))
```

Output in the next slide >>

Visualizing selected regions

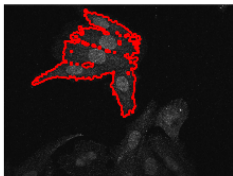
Merge



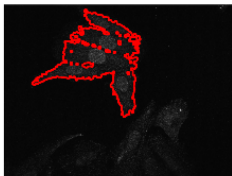
Pixel Set



Channel One



Channel Two



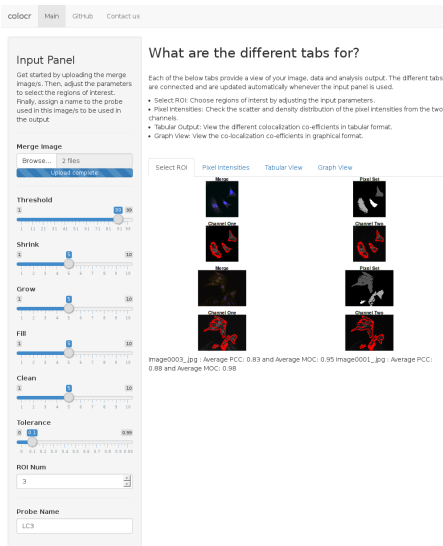
<< Code in the previous slide

The interactive web app

The package provides a simple web app to interactively determine the selection parameters and apply the rest of the workflow. To launch the app, run the following

```
# run the shiny app
colocr_app()
```

Or visit this page.



Selecting multiple regions

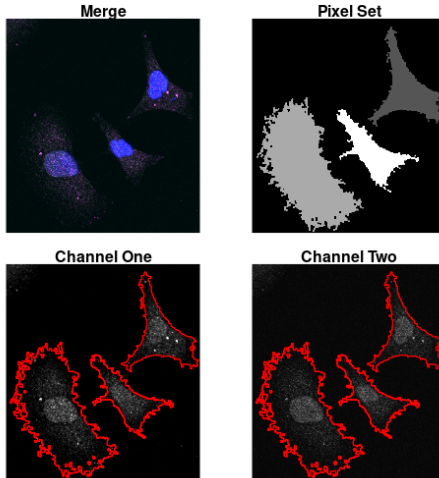
If we want `roi_select` to select multiple regions, we modify the argument `n` and the selection parameters.

- `threshold`: as before
- `shrink`: fitting pixels in a structure
- `clear`: removing small isolated elements
- `n`: the number of regions to select

```
# load 'image_2.png' as img2  
img2 <- image_load('image_2.png')  
  
# select and show three regions of interest  
img2_roi <- roi_select(img2, threshold = 90,  
  shrink = 10, clean = 10, n = 3)  
roi_show(img2_roi)
```

Output in the next slide >>

Selecting multiple regions



<< Code in the previous slide

Pearson's correlation coefficient

Pearson's correlation coefficient (PCC) is the co-variance of the pixel intensity between two channels. The mean of the intensities is subtracted from each pixel which makes the coefficient independent of the background level.

The PCC is calculated as follows

$$\text{PCC} = \frac{\sum_i (R_i - \bar{R}) \times (G_i - \bar{G})}{\sqrt{\sum_i (R_i - \bar{R})^2 \times \sum_i (G_i - \bar{G})^2}}$$

where R_i and the G_i are the intensities of the red and green channels and the \bar{R} and \bar{G} are the average intensities.

PCC value of 1 means positive and -1 negative perfect correlation. 0 means there is no correlation between the pixel intensities.

Manders Overlap Coefficient

Manders Overlap Coefficient (MOC) does not require subtraction of the mean. Therefore, the values are always between 0 and 1. Also, the MOC is independent of signal proportionality.

$$\text{MOC} = \frac{\sum_i (R_i \times G_i)}{\sqrt{\sum_i R_i^2 \times \sum_i G_i^2}}$$

where R_i and the G_i are the intensities of the red and green channels.

Calculating correlation statistics

The `colocr` package implements both PCC and SCC in `roi_test`.

Invoking the test is a one-function call on the selected regions of interest.

```
# Calculate the PCC  
tst <- roi_test(img2_rois, type = 'both')  
tst
```

	pcc	moc
1	0.8798341	0.9303114
2	0.8963802	0.8990838
3	0.8882729	0.9386888

`roi_test` returns a data.frame with a column for each of the desired statistics: 'pcc' and 'moc', and a row for each region of interest.

References

1. Ahmed M, Lai TH, Kim DR. colocr: an R package for conducting co-localization analysis on fluorescence microscopy images. PeerJ. 2019;7:e7255. doi:10.7717/peerj.7255.
1. Dunn KW, Kamocka MM, McDonald JH. A practical guide to evaluating colocalization in biological microscopy. Am J Physiol Cell Physiol. 2011;300(4):C723-C742. doi:10.1152/ajpcell.00462.2010.

Summary

What you've learned

- Fluorescence image analysis
- Understanding images
- Loading images in R
- Selecting regions of interest
- Calculating correlation statistics

What's next

- Practice ([Link](#))
- Homework ([Link](#))
- What to learn next
 - Introduction to R (Data Camp Free Interactive Course)
 - Intermediate R Course (Data Camp Free Interactive Course)
 - R for Data Science (Book)