

Assignment C: Analyzing treatment effect on breast cancer cell line using FISH

Course : Image Processing and Quantitative Data Analysis

Instructor : Marten Postma

Teaching assistants : Aaron Lin, Aoming Sun, Catherine Chia

Date: 12th June, 2023

Table of content

1.0 Introduction	1
2.0 Tasks and questions (Total 10 points)	3
2.1 Image data inspection (1 point)	3
2.2 Image processing workflow (4 points)	3
2.3 Data analysis and discussion (5 point)	3
3.0 Reference	3

1.0 Introduction

MCF-7 is a human breast cancer cell line commonly used in studying drug efficacy. Using Fluorescent in situ hybridization (FISH) technique, a cancer marker in MCF-7 cell line such as *ARAG* gene can be tagged and its transcription can be quantified numerically and spatially. In this assignment, the experiment treated the MCF-7 cell lines with one of these 3 drugs:

- DMSO (control)
- JQ1
- TSA

Both JQ1 and TSA are inhibitors of the *ARAG* gene and the expression of *ARAG* can be used as a proxy to the strength of these inhibitors. The goal of this assignment is to compare the effect of drug treatment by describing the mRNA counts within the nucleus or in the cytoplasm of each cell in each condition.

Image data in folders: input/DMSO, input/JQ1, and input/TSA

The 3D image data (Figure A) in these folders are individual channels (DAPI for nuclear marker, single-molecule (sm) FISH for *ARAG* mRNA) for 3 conditions: DMSO (control group), JQ1 (drug treatment) and TSA (another drug treatment). The Z-axis represents the thickness of the cell and there is one cell for each condition.

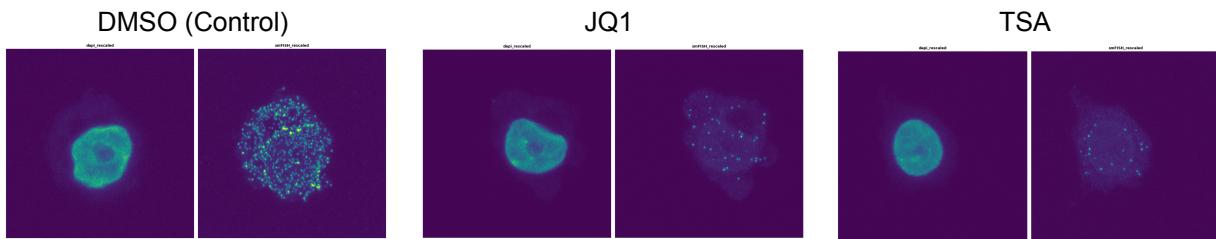


Figure A Datasets for all conditions (DMSO, JQ1 and TSA). Left image of each condition indicates DAPI labeling for the nucleus, and the right images indicate the FISH-tagged ARAG mRNA within the cell.

Additional image data in folders: output/DMSO, output/JQ1, and output/TSA

To facilitate the learning process, we have generated the segmentation masks for each condition, using the notebook provided by the authors of BigFISH (Imbert and Mueller, 2022). The cell mask (*_cell_label.tif) describes the region of the cell, whereas the nuclear mask (*_nuc_label.tif) indicates the region of the nucleus.

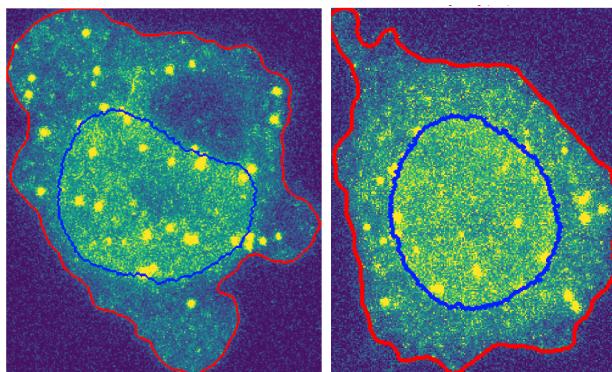


Figure B Segmentation masks projected on the cell (red border) and nucleus (blue border). The spots are FISH-tagged ARAG mRNA. MCF-7 cell perturbed by JQ1 (left) and TSA (right).

General instruction

This assignment focuses on the understanding of the workflow. Thus there is minimal coding effort required in this assignment, but you are required to create a Conda environment and run the Jupyter Notebooks with the provided data. Also, we would like you to explore the documentation of BigFISH (<https://big-fish.readthedocs.io/en/stable/>) to answer the assignment questions.

To begin, go to <https://github.com/catherinechia/big-fish-IPQDA> and follow the installation instructions on the page.

There are several special instructions in the Jupyter Notebooks:

1. The cell/chunk following this heading will require you to uncomment/comment the correct file path. You should use the images from one condition to run the scripts (i.e. use DMSO images from the start till the end of both Notebooks)

(IPQDA: Change values here)

2. The plot or table following this type of heading will be used in answering the questions.

(IPQDA: Figure 1 - Spot detection)

2.0 Tasks and questions (Total 10 points)

2.1 Image data inspection (1 point)

- Question 1: What could be the reason that the FISH-tagged ARAG mRNA is not confined to the nucleus?

2.2 Image processing workflow (4 points)

- Question 2: Explain the difference between **Figure 1 and 2** in the notebook “5 - Detect spots.ipynb” by elaborating the function `detection.decompose_dense()`.
- Question 3: What is the difference between spots and clusters (**Figure 3** in notebook “5 - Detect spots.ipynb”)? From a biology perspective, why does it matter to find the clusters?

2.3 Data analysis and discussion (5 point)

- Question 4: Combine **Figure 4** and **Table 1** (a pair for each condition) in notebook “6 - Extract cell level results.ipynb” in a presentable way and describe the results.
- Question 5: Are the mRNA counts in the nucleus and in the cytoplasm maintained around the same proportion in each condition? If not, how would you describe the drug efficacy with respect to the spatial information of the mRNA counts?
- Question 6: Write a conclusion about which drug has a greater treatment effect (based on your answer for Question 5)?
- Question 7: Do you think making a conclusion based on one cell per condition is an ideal approach? Why and why not?

3.0 Reference

Arthur Imbert, Wei Ouyang, Adham Safieddine, Emeline Coleno, Christophe Zimmer, Edouard Bertrand, Thomas Walter, Florian Mueller. FISH-quant v2:a scalable and modular analysis tool for smFISH image analysis. bioRxiv (2021) <https://doi.org/10.1101/2021.07.20.453024>

Imbert A. and Mueller, F. (2022) 4 - Segment nuclei and cells.ipynb

<https://github.com/fish-quant/big-fish-examples/blob/master/notebooks/4%20-%20Segment%20nuclei%20and%20cells.ipynb>