**Uncovering dynamic transcriptional regulation of methanogenesis via** **single-cell imaging of archaeal gene expression**

**1. Introduction**

This code package is developed for analyzing confocal image stacks of microbial cells labeled using smHCR-FISH. In this document, we provide essential guidance for setting up and using the code.

The code requires pre-installation of MATLAB (2023a) with the image analysis, curve fitting, optimization, and parallel processing toolboxes, and the Bio-Formats package (available at <https://www.openmicroscopy.org/bio-formats/downloads/>). ImageJ (v1.54j) also needs to be installed for cell segmentation. For optimal performance, we recommend a 64-bit computer system with a multi-core processor (≥36 cores) and large RAM (≥256 GB). To install the package, simply copy the “smHCR-FISH” folder to your computer and add “smHCR-FISH/MATLAB” to the MATLAB search path.

**2. Image conversion and pre-analysis processing**

First, navigate to “smHCR-FISH/Confocal/” and create a subfolder (e.g., “Exp1/”) for each confocal image file, which is downloadable from http://gofile.me/4yuzx/zbFlbhcpr/LMO-1 FISH data. In addtion, download the folder “Calibration” into the “smHCR-FISH” folder. Note that each confocal file contains multiple fields of view (FOVs) from different growth time points of a given round of smHCR-FISH imaging, while results from distinct rounds of a multi-round smHCR-FISH experiment are stored as separate files. For demonstration, you may download the file for a single-round experiment: “lmo\_mcrA\_b3\_mtrA\_b5\_mch\_b2\_b2\_2\_2\_2\_2\_2024\_09\_15\_\_10\_41\_41-Airyscan Processing-03.czi”.

Next, open the Excel file “smHCR-FISH/Duallist.xls”, which stores the information about image files for processing. Enter your data in the following format:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Folder name** | **Leave Blank** | **Channel index** | **Tiling method** | **Channel name** | **Running index** |
| Confocal/Exp1/ |  | [4,4,3,2,1] | [1,1] | {'A647','TMR','A488','DAPI'} | T |

Here, the four imaging channels (Alexa 647, TMR, Alexa 488, and DAPI) correspond to three RNA signals and the DNA signal, respectively. After saving, run the following scripts in MATLAB:

1. To convert the confocal CZI files into TIF images, run “cziconvert2('Duallist.xls');”. This creates a “stacks/” subfolder within your data folder (e.g., “Confocal/Exp1/”) containing all TIFF images extracted from the raw image file. Specifically, images from different FOVs are stored in separate subfolders inside “stacks/”.
2. With in“stacks/”, find an Excel file “matchlist.xls”, which stores the information about each FOV subfolder. Open this file and copy the first column content to the third column.
3. For multi-round experiment, run “align\_stacknew('Duallist.xls');” to align images from different rounds. The result of this alignment is saved as a file “Rectify\_pair.fig” in each FOV subfolder. To manually refine alignment results, run “manual\_align”.

\* Skip this step for single-round experiment data.

1. To extract basic information from images for further analysis, run “Dualprocess1;”. This creates a subfolder “Results/” containing pre-analysis results.

**3. Cell segmentation**

1. Perform automatic segmentation using the ‘StarDist’ Plugin in ImageJ, then manually refine the regions of interest (ROIs). Create a subfolder “RoiSet/” to save the manual refinement results and create an Excel file named “Reg\_list.xls” to record the file name of each segmentation result.
2. Create a subfolder named “masks/”, copy “stacks/matchlist.xls” into it, and run “mask\_rec\_single” to convert segmentation results into a format compatible with the MATLAB code. The results are stored in “masks/”.

**4. smHCR-FISH spot analysis**

To identify smFISH spots and quantify their intensities, run the following scripts:

1. For automatic spot quantification, use the MATLAB GUI “stack\_RNA\_01\_new”. Set the “Analysis types” to “foci 1”, “foci 2”, or “foci 3” for TMR, Alexa 647, or Alexa 488 channels, respectively, and click “Auto” to initiate RNA spot identification. This creates a subfolder named “Histogram”, “Histogram\_RNA2”, or “Histogram\_RNA3” containing mRNA spot identification results for the corresponding channel. Results for each FOV are stored as a separate Excel file with “\_raw.xls” suffix.
2. To review and refine automatic identification, use the MATLAB GUI “stack\_RNA\_check\_01”. Set the “RNA Channel” to “ch1”, “ch2”, or “ch3” for TMR, Alexa 647, or Alexa 488 channels, respectively.
3. To extract the single-mRNA intensity for a given mRNA species, run the following command: “hist\_fit\_RNA\_stage(‘Duallist.xls’,FOV\_index,hist\_name);”. Here, the input parameter “FOV\_index” refers to indices of stationary-phase FOVs (correspond to row numbers in “matchlist.xls”) for analysis, while “hist\_name” specifies the relevant spot-identification subfolder, i.e., “Histogram”, “Histogram\_RNA2”, or “Histogram\_RNA3”. For example, to extract the single-mRNA intensity for the TMR channel from the first three FOVs of the image file, you may run ‘hist\_fit\_RNA\_stage('Duallist.xls',1:3,'Histogram/');’.

**5. Analysis of transcriptional regulation**

1. With cell segmentation and spot identification complete, quantify gene expression in individual cells using “foci\_mask\_profile('Duallist.xls',ch\_index,hist\_name);”. Here, “ch\_index” refers to the fluorescent channel for processing, i.e., “1” (default) for Alexa 647, “2” for TMR, and “3” for Alexa 488, while “hist\_name” specifies the relevant spot-identification subfolder, i.e., “Histogram”, “Histogram\_RNA2”, or “Histogram\_RNA3”. Results of this analysis are stored as Excel files (one for each FOV) with “N\_Ifoci.xlsx” suffix in the corresponding spot-identification subfolder.
2. Integrate each gene’s expression data (across multiple FOVs and growth time points) into an Excel file, with data from each growth time point stored in each columns.
3. Run “TwoState\_FSP\_NSX\_st\_twocopy\_P0\_simu” to fit the mRNA distributions to a two-state model.