Quantification of smFISH images

This quantification pipeline is divided into 5 parts:

- 1st part: oocytes outlines definition and extraction.
- 2nd part: image preprocessing.
- 3rd part: condensates segmentation.
- 4th part: generation of outlines and filtered images for smFISH spots detection in the cytosol.
- 5th part: smFISH spots detection and filtering.

Requirements:

- MATLAB v R2018b or higher
  - Required toolboxes:
    - Image Processing Toolbox
    - Statistics and Machine Learning Toolbox
    - Parallel Computing Toolbox
    - Curve Fitting Toolbox
    - Optimization Toolbox
  - Additional requirements:
    - FISH-quant[1,2,#]

    [#] https://bitbucket.org/muellerflorian/fish_quant/src/master/

Set up:

- Copy the scripts provided (~/scripts_smFISH_quantification) into the MATLAB folder (Documents directory).
- Install FISH-quant:
  a. Go to ***https://bitbucket.org/muellerflorian/fish_quant/src/master/*** [1,2].
  b. Click Downloads and then download the repository[1,2].
  c. Unzip the file in the MATLAB folder (Documents directory)[1,2].
  d. Go to ~/Documents/MATLAB/FISH_quant/Documentation[1,2].
  e. Open **FISH-QUANT__Tutorials.pdf**[1,2].
  f. Follow the section **1.1 Install FISH-quant for Matlab**[1,2].

Important:
How does this pipeline work?

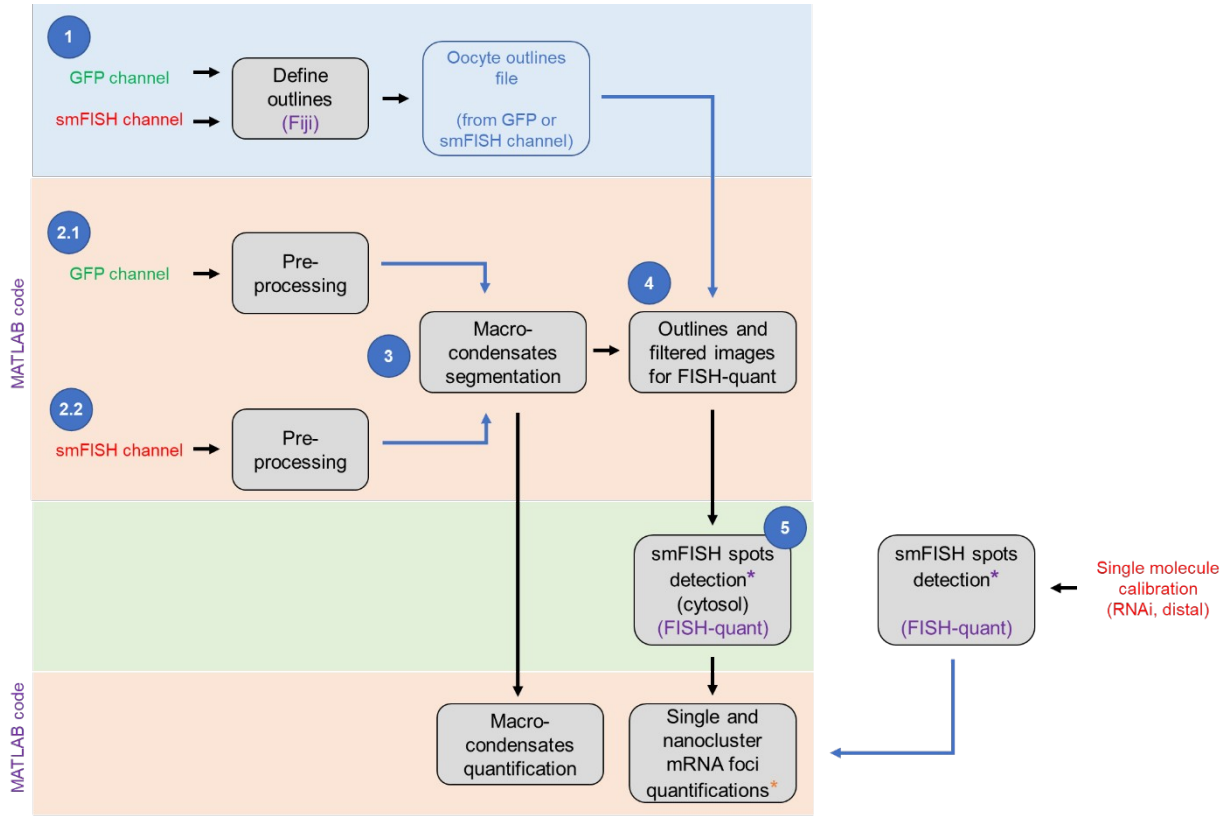


Figure 1. smFISH image quantification pipeline. It uses Fiji, FISH-quant, and custom MATLAB code. (*) These steps can be done in batch. For the rest of steps several MATLAB sections can be open simultaneously.

Provided image example

The image data is provided at ~/example image

It consists of two channels (already separated as independent images):

- GFP channel: “w4_GFP.tif”
- smFISH channel: “w4_Spn4.tif”

An example oocyte outlines file (see Fig. 1) is also provided:

- “w4_GFP.zip”

1st part: oocyte outlines definition and extraction

Requirements:

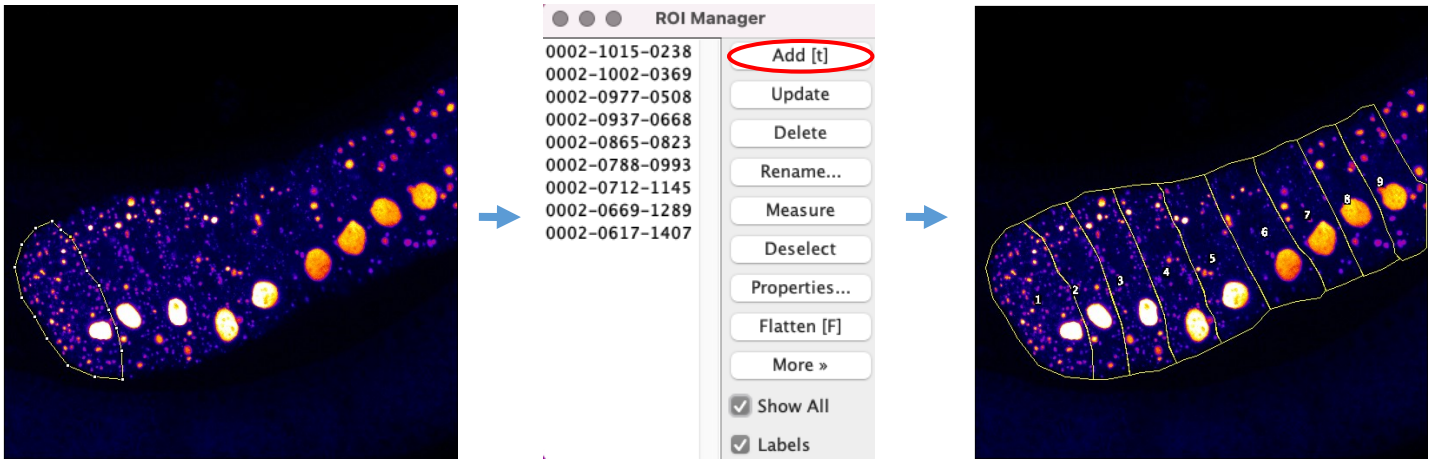
- Fiji (Image J)
- R
- Optional: RStudio

Outlines determination

Procedure:

1. Open Fiji (ImageJ).
2. Open *RIO Manager*
   - Analyze > Tools > RIO Manger…
3. Using Fiji, open the image you want to analyze. Here you can use either the GFP or smFISH channel (see section Provided image example and Fig.1).
4. Using the "polygon selections" tool manually draw outlines that delimit the oocyte borders.

   - Each oocyte outline can be treated as an individual Region Of Interest (ROI).
   - To delimit oocytes, a max projection could be useful as gonad dimensions are variable along the *z*-plane.
   - Each oocyte outline (ROI) can be saved by clicking on the "Add [t]" bottom located in the *RIO Manager* window, as shown below:



5. Once all outlines are added, select them and save them as a "*.zip*" file. **Important:** name the file with the same name as the image where the outlines are coming from.
   - In the case of only one outline, it will be saved as a "*.roi*" file.

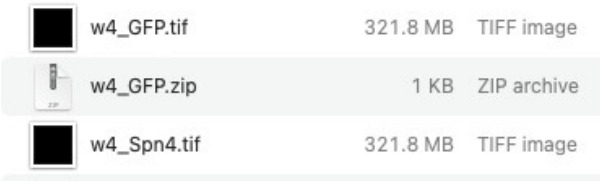
Outlines extraction of xy coordinates from outlines (required for 4rd part)
Example data and description:
Example images and outline file for this guide can be found in the folder ~/example_image

Images were obtained imaging a smFISH experiment. It consists of two channels (already separated as independent images):

- GFP channel: "w4_GFP.tif"
- smFISH channel: "w4_Spn4.tif"

An example oocyte outlines file (see Fig. 1) is also provided. You can obtain it as explained in the previous section (Outlines determination)

- "w4_GFP.zip"



Procedure:

6. Data should be organized in a folder as shown just above:
   For each *"name".zip* outline file, a matching *"name".tif* image must be in the folder.
   - For this example, only one outline file will be processed, alternatively several outline files with their corresponding images can be processed simultaneously.

7. Using Fiji, open the script "*BATCH_improved_outlines_oocyte_extraction.ijm*"
   Script location at ~/image_outlines

8. Run it and select the folder containing the images and oocyte outlines (ROIs) to analyze.
   This script extracts the xy coordinates from the ROIs delimited and saved using Fiji.

   Output: *"name"_Cells.csv* files with xy coordinates of ROIs.

9. Using R or RStudio, open the script "*BATCH_automatic_outlines_reorganization_by_columns.R*
   Script location at ~/image_outlines

10. Run the script and select any file within the folder containing the images and ROIs to analyze.
    This script reorganizes the outlines files, so they are readable in the image processing step.

    Output: *MOD_"name"_Cells.csv* files with reorganized xy coordinates of ROIs.

Optionally, you can do this reorganization manually as shown below:

	A	B
1	X_coodinate_cell_0	
2	747	
3	798	
4	840	
5	932	
6	1036	
7	1152	
8	1188	
9	1238	
10	1261	
11	.	
12	.	
13	.	
14	666	
15	708	
16	y_coodinate_cell_0	
17	1140	
18	1167	
19	1197	
20	1244	
21	1276	
22	1284	
23	1340	
24	1332	
25	1337	
26	1354	

	A	B	C	D	E	F	G	H	I
1	X_coodinat	y_coodinat	X_coodinat	y_coodinat	X_coodinat	y_coodinat	X_coodinat	y_coodinate_cell_3	
2	747	1140	242	535	150	222	234	12	
3	798	1167	300	559	273	300	411	78	
4	840	1197	369	603	444	357	561	132	
5	932	1244	441	648	627	420	741	180	
6	1036	1276	540	690	777	513	993	285	
7	1152	1284	612	747	990	594	1104	366	
8	1188	1340	765	810	1131	663	1353	438	
9	1238	1332	912	879	1290	720	1512	498	
10	1261	1337	1035	918	1467	786	1512	678	
11	1317	1354	1206	984	1509	831	1512	834	
12	1289	1388	1275	1014	1509	1005	1464	786	
13	1224	1479	1326	1080	1461	1125	1290	723	
14	1189	1514	1468	1116	1455	1113	1140	669	
15	505	1513	1328	1356	1329	1083	1011	603	
16	442	1498	1242	1332	1281	1020	885	552	
17	349	1468	1188	1338	1083	936	777	516	
18	196	1456	1152	1281	939	885	615	414	
19	85	1369	1044	1272	777	816	435	354	
20	1	1237	930	1245	624	750	291	309	
21	3	1029	849	1200	534	690	156	228	
22	3	903	753	1140	393	624	159	180	
23	7	751	717	1143	291	555	192	87	
24	85	808	594	1095	171	501			
25	136	826	525	1065	63	438			
26	178	850	477	1029	105	330			
27	226	877	381	996					
28	286	910	288	912					

How to modify "*BATCH_improved_outlines_oocyte_extraction.ijm"*

Line 1 >

//... Modify if needed (File Identifiers)...

Line 2 >

file_ext = "zip"; *// ROI file extension*

Line…

Only line 2 should be modified. If needed replace "*zip*" by "*roi*".

2nd part: image preprocessing.

Procedure (for each channel, GFP and smFISH) do the following:

11. Open MATLAB

12. In the command window type:

- *edit smFISH_Preprocessing_Script*

This file will open:

```
%%  ----------------- smFISH preprocessing
% Description of the stack
% There are two possibilities:
% = 1 if stack has beginning and ending frames out of focus (whole gonad)
% = 0 if stack is in focus (usually a z-section within the gonad)
is_stack_out_of_focus = 1; %
smFISH_channel = 0; % type of image: if FISH  = 1, if GFP = 0

% Parameters
quantile_BGD_subtraction = 0.9; % quantile to subtract BGD value (initial filter-
ing), 0 if it's unknown
% if quantile_BGD_subtraction = 0, two BGD approximations are computed:
different_BGD_approximations = 1; % 1 for different functions, 2 for same function
with different mask sizes

% image ranking
block_size_for_image_ranking = 2; % faster results with larger values like 4, 8,
16; but less detail

% run
smFISH_image_Preprocessing
```

13. If needed, modify the parameters for preprocesssing, then save changes:

13.0.1. Line # 6: `is_stack_out_of_focus`

This scrip is optimized for two options, either a complete gonad in the z-plane with out-of-focus planes (`is_stack_out_of_focus = 1`), or a section z-section with NO out-of-focus planes (`is_stack_out_of_focus = 0`).

13.0.2. Line # 7: `smFISH_channel`

Make `smFISH_channel = 1` for smFISH channel
Make `smFISH_channel = 0` for GFP channel

13.0.3. Line # 10: `quantile_BGD_subtraction`

`quantile_BGD_subtraction = 0.9` is recommended for most cases.
To select a value during the preprocessing make `quantile_BGD_subtraction = 0`

13.0.4. Line # 15: `block_size_for_image_ranking = 2;`

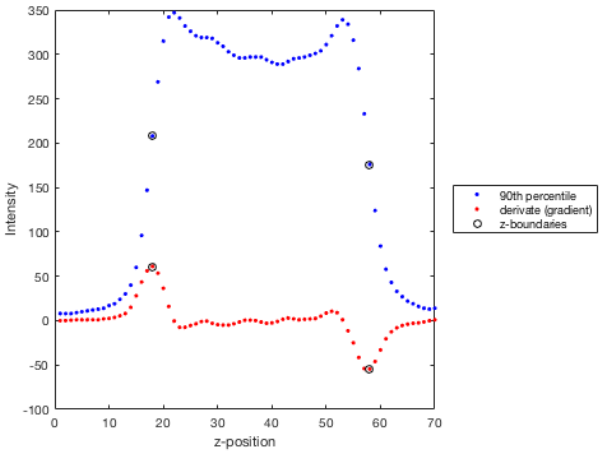
```
faster results with larger values like 4, 8, 16; but less detail
```

14. In the command window type: *smFISH_Preprocessing_Script*

15. The script will ask you to select either a GFP or smFISH image.

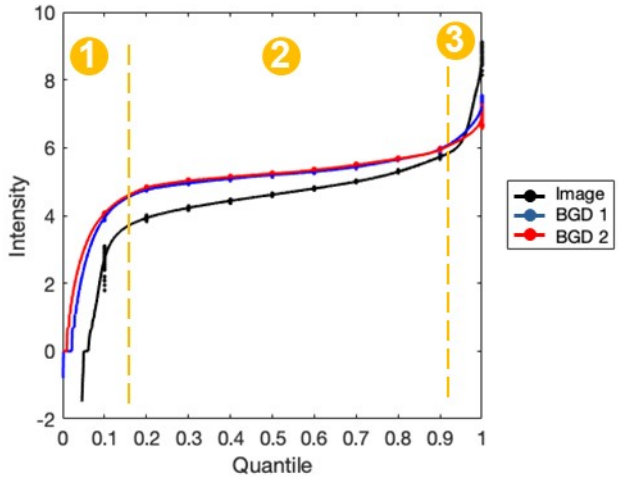
15.0.1. The preprocessing steps are the same, the distinction is only used to name the outputs properly.

16. If `is_stack_out_of_focus = 1`, The script determines the oocyte z-boundaries of the stack (oocytes z-edges), as follows:



*** This method won't work if the z-stack acquisition exhibits bleaching in the z-plane or if you if it doesn't have out-of-focus frames both at the beginning and end of the z-stack.

17. Then:

If `quantile_BGD_subtraction = 0.9`	if `quantile_BGD_subtraction = 0`
• The script will subtract the background. • Image ranking will be performed. • And output will be saved.	• The script will perform two background approximations (BGD1 and BGD2):  The curve displays three regions: 1. slide background, 2. oocyte background. 3. brighter objects like smFISH clusters. `quantile_BGD_subtraction` should be the quantile that separates region 2 from region 3.

	By default, `quantile_BGD_subtraction = 0.9` • Type in the command window the quantile you want. • The script will subtract the background. • Image ranking will be performed. • And output will be saved.

By channel the script will save a MATLAB workspace. For instance:

- WS_*name-smFISH-cahnnel*_FISH.mat
- WS_*name-GFP-cahnnel*_GFP.mat

3rd part: condensates segmentation

Example data and description

For this step, the following files are required:

location: ~/example_image

- GFP channel image
- smFISH channel mage

location: ~/example_image/output_image_preprocessing (2nd part)

- WS_*name-smFISH-cahnnel*_FISH.mat
- WS_*name-GFP-cahnnel*_GFP.mat

location: ~/example_image/output_outlines_reorganization (1st part)

- *MOD_"name-channel"_Cells.csv*

For this example, the, the following files are provided:

- w4_**Spn4**.tif
- w4_**GFP**.tif
- WS_w4_Spn4_**FISH**.mat
- WS_w4_GFP_**GFP**.mat
- MOD_w4_GFP_**Cells**.csv

Procedure:

18. Make a folder with the 5 elements described in “example data and description”.

19. Type in the command window:

edit smFISH_Segmentation_Outlines_Script

20. The script below will appear (If needed, modify accordingly to instructions). Then save changes.

```
%  ----------------- smFISH image analysis (segmentation)
%% File ID indicators
files = struct; % do not modify
files.FISH_img = 'Spn4'; % file indicator for smFISH image
files.FISH_ws  = 'FISH'; % file indicator for smFISH workspace
files.GFP      = 'GFP';  % file indicator for GFP image and workspace

%% Microscope parameters
% Define pixel size in nm
pixel_size = struct; % do not modify
pixel_size.xy = 49;  % pixel size in x and y
pixel_size.z  = 250; % pixel size in z

% Acquisition parameters
s = struct; % do not modify
s.Em    = 568;        % cy3 Emission
s.Ex    = 554;        % cy3 Excitation
s.NA    = 1.4;        % Numerical aperture
s.RI    = 1.518;      % Diffractive index
s.type  = 'confocal'; % Microscope type
s.Pixel = pixel_size;

%% Segmentation parameters
options = struct; % do not modify
options.Int_Threshold_GFP  = 0.6; % Normalized [0 - 1] Intensity Threshold: GFP
options.Int_Threshold_FISH = 0.8; % Normalized [0 - 1] Intensity Threshold: smFISH
options.Volume_Threshold   = 32;  % Volume threshold for segmentation (in pixels)
options.Dilate_Granule     = 1;   % if 1, condensate are dilated to include edges

%% Outlines settings
options.FQ_outline       = 1; % if 1, outlines for FQ are an output
options.Exclude_Granules = 1; % if 1, creates images and outlines without granules
options.Save_Coordinates = 1; % if 1, segmented objects coordinates for are saved

%% Alternative masking
Filt_options = struct;  % do not modify
Filt_options.gauss2x  = 1;
Filt_options.bgd_xy   = 5;
Filt_options.bgd_z    = 5;
Filt_options.spots_xy = 0.1;
Filt_options.spots_z  = 0.1;
Filt_options.output   = 0;
options.Filt          = Filt_options;

%% RUN
Segmentation_Outlines_run
```

20.0.1. The following values should be adapted to different experimental conditions, for example; a different mRNA target, type of microscopy, among others. However, should be maintained if you are analyzing the same experiment.

- `options.Int_Threshold_GFP  = 0.6;`
  % Normalized intensity threshold for segmentation of the GFP image.
  Values accepted from 0 to 1. Values close to 0 would be background while values close to 1 will be condensates
- `options.Int_Threshold_FISH = 0.8;`
  % Normalized intensity threshold for segmentation of the smFISH image.
  Values accepted from 0 to 1. Values close to 0 would be background while values close to 1 will be condensates
- `options.Volume_Threshold   = 32;`
  % Volume threshold for segmentation (in pixels)
- `options.Dilate_Granule     = 1;`
  % if = 1, condensates are dilated to make sure edges are included. This is a very important steps as edges not segmented might generate false positives while detecting cytosolic spots. IMPORTANT: Dilation needs to be done by condensate, so this step might take time depending on computer specifications.

20.0.2. If you need to make test without running the whole script (specially the dilation that takes most of the time), in step 18 type in the command window:

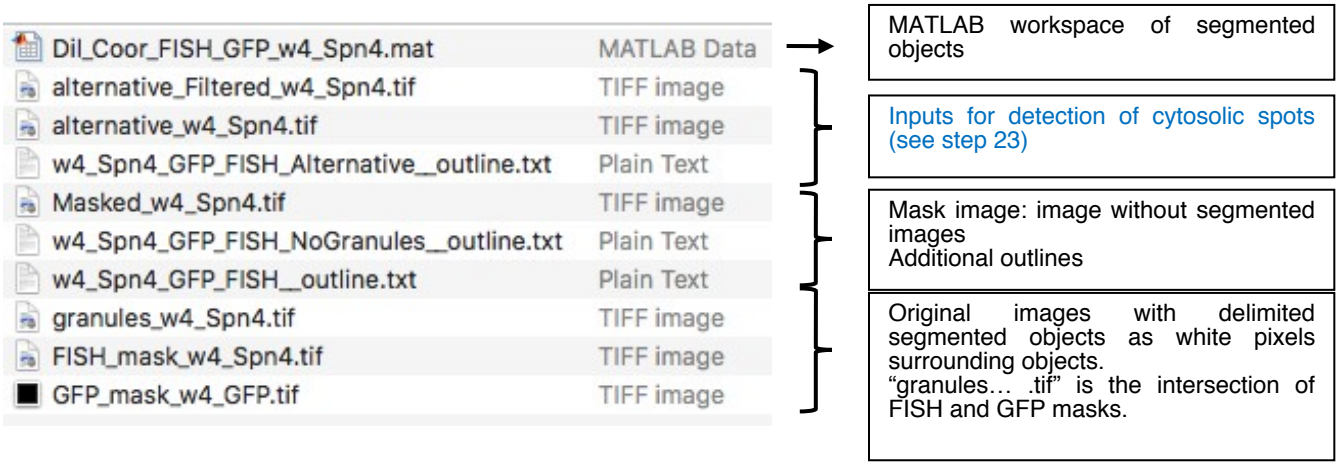
edit smFISH_Segmentation_Script_TEST

21. Accordingly, type in the command window:

smFISH_Segmentation_Outlines_Script
or
edit smFISH_Segmentation_Script_TEST

22. Select the folder created in step 18.

Segmentation outputs



4th and 5th part:
Generation of outlines and filtered images for smFISH spots detection in the cytosol, smFISH spots detection and filtering.

The segmentation script also generates the outlines, raw image, and filter image required to run FISH-quant[1,2,3] to detect smFISH spots. For more information, refer to the already published manuals[1,2,3].

23. In order to detect cytosolic smFISH spots after running this segmentation (Fig. 1), please use the following files as FISH-quant inputs (outputs from 3rd part: condensates segmentation):
- alternative… .tif image
- alternative_Filtered… .tif
- …Alternative__outline.txt

24. Once in FISH-quant[1,2,3], single or batch analysis modes, do not perform the filtering step.
If you want to modify the filtering parameters go to step **20** and make changes as follows:
- ```
  %% Alternative masking
  Filt_options = struct;  % do not modify
  Filt_options.gauss2x    = 1;
  Filt_options.bgd_xy     = 5;   % Kernel BGD [pixel]: XY (see FISH-quant)
  Filt_options.bgd_z      = 5;   % Kernel BGD [pixel]: Z  (see FISH-quant)
  Filt_options.spots_xy   = 0.1; % Kernel SNR [pixel]: XY (see FISH-quant)
  Filt_options.spots_z    = 0.1; % Kernel SNR [pixel]: Z  (see FISH-quant)
  Filt_options.output     = 0;
  options.Filt             = Filt_options
  ```

25. Use FISH-quant as previously described[1,2,3]
- Recommendations
  - Set analysis parameters with the original images
  - Create your file with settings (see FISH-quant documentation[1,2,3])
  - Use batch mode to run the images (step 23)
  - Select Use filtered images

Integration with FISH-quant

Condensates quantification (Fig.1, see Supplementary Information[4]) is performed after the detection of cytosolic smFISH spots.

To avoid any detection of spots within condensates by FISH-quant, the output images (step 22 and 23) are labeled, so:

- The only possible detection of false positive smFISH sports occurs at condensates edges.
- After detection in FISH-quant, false positive spots at the condensate edges can be filtered out because they would have a unique and very high “Pixel intensity (Raw)” value that: (1) it is not preset in the image and (2) it is biased to the right side as shown in the following example:
  ```

Before filtering out false positives from condensates edges

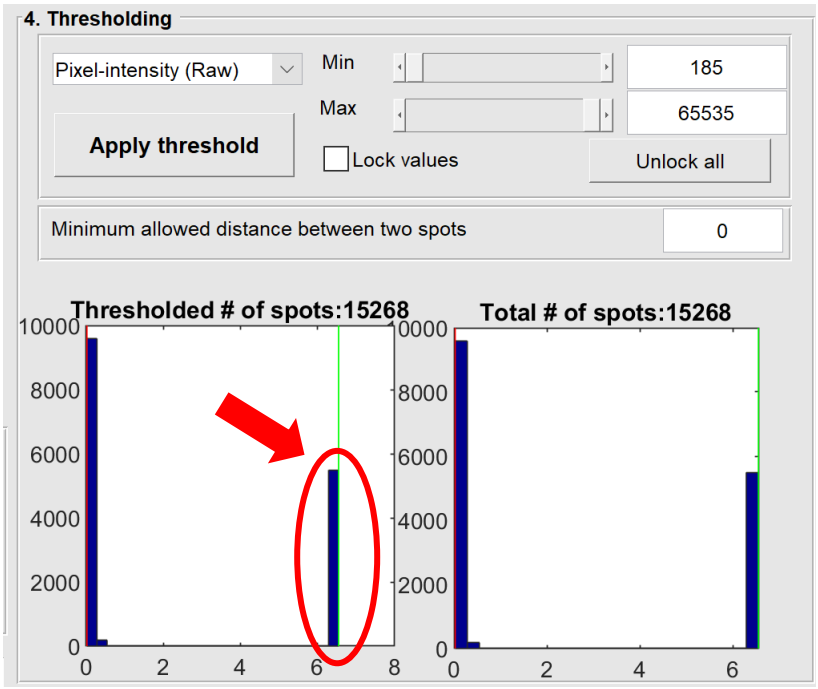


Image from FISH-quant v3 main interface[1,2,3]

After filtering out false positives from condensates edges

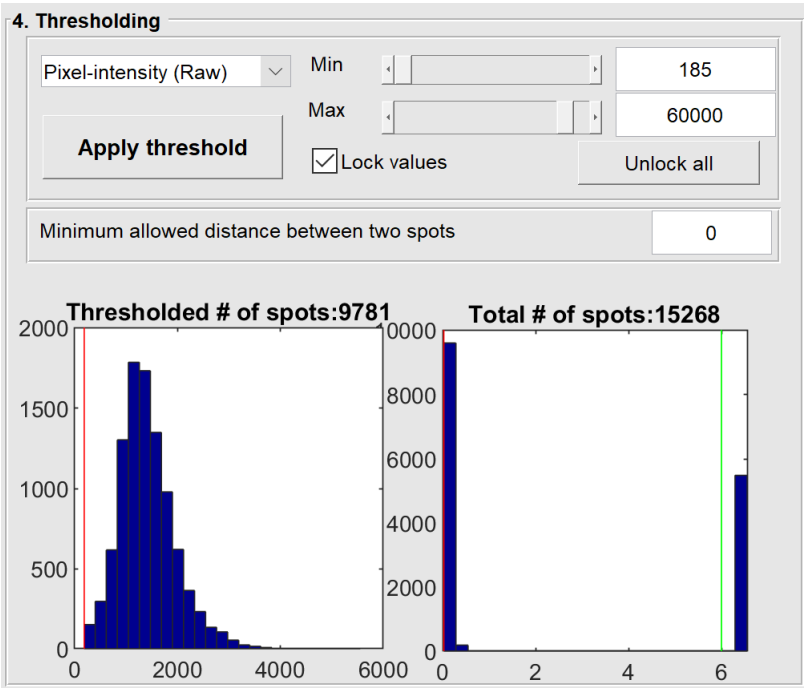


Image from FISH-quant v3 main interface[1,2,3]

References

1. Mueller, F., Senecal, A., Tantale, K. *et al.* FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat Methods* **10**, 277–278 (2013). https://doi.org/10.1038/nmeth.2406

2. Tsanov, N., Samacoits, A., Chouaib, R. *et al*. smiFISH and FISH-quant – a flexible single RNA detection approach with super-resolution capability, *Nucleic Acids Research* 44 (22), e165 (2016). https://doi.org/10.1093/nar/gkw784

3. https://bitbucket.org/muellerflorian/fish_quant/src/master/

4. Cardona, AH., Ecsedi, S. *et al*. Robust sorting and buffering within condensates control transcriptome stoichiometries (2023). Submitted