

FISH-quant: tutorials

In this tutorial we explain the basic workflow of the analysis of FISH images in FISH-quant based on the example data-set. You can familiarize yourself with the main functionality of FISH-quant by following this tutorial. We will refer several times specifically to the example data-set and invite the user to change different parameters and investigate the effect on the analysis. These examples are indicated with boxes like the following:

Example: load image BGH_03_CY3.tif

We provide screen-shot of the most important steps explained. For a more **detailed explanation of the different options** and advanced features please consult the detailed user manual.


More details and links for **example data** and a **more detailed user manual** can be found at https://bitbucket.org/muellerflorian/fish_quant

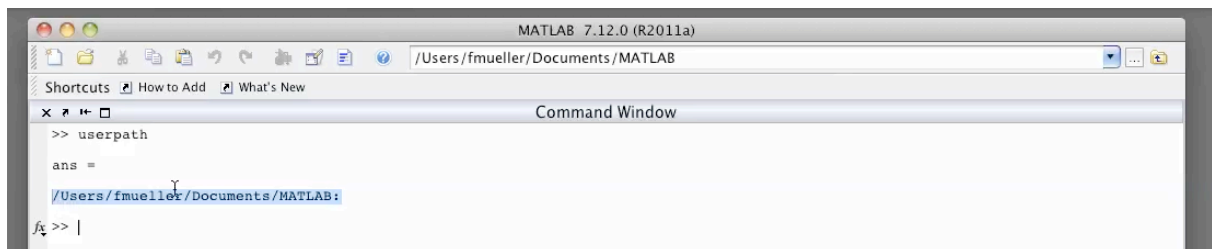
| | |
|--|-----------|
| 1. Getting started | 2 |
| 1.1. Install FISH-quant for Matlab | 2 |
| 1.2. Typical files produced in FISH-quant - overview of example data set..... | 3 |
| 1.3. Define a folder-structure for your project | 3 |
| 1.4. Inspect example data | 4 |
| 2. Defining outline files manually | 6 |
| 2.1. Define outlines | 6 |
| 2.2. [Optional] Automatic detection of nuclei | 8 |
| 2.3. [Optional] Automatic detection of transcription sites..... | 8 |
| 3. Defining outline files automatically | 10 |
| 3.1. 2D Projections | 10 |
| 3.2. Cell segmentation | 11 |
| 3.3. Automated outline generation..... | 12 |
| 4. Mature mRNA counting..... | 13 |
| 4.1. Mature mRNA detection on individual images..... | 13 |
| 4.2. Mature mRNA detection in batch mode | 19 |
| 5. Nascent mRNA counting | 22 |
| 5.1. How to obtain the necessary files for the quantification | 22 |
| 5.2. Transcription site quantification with integrated intensity..... | 23 |
| 5.3. Transcription site quantification in batch mode | 24 |

1. Getting started

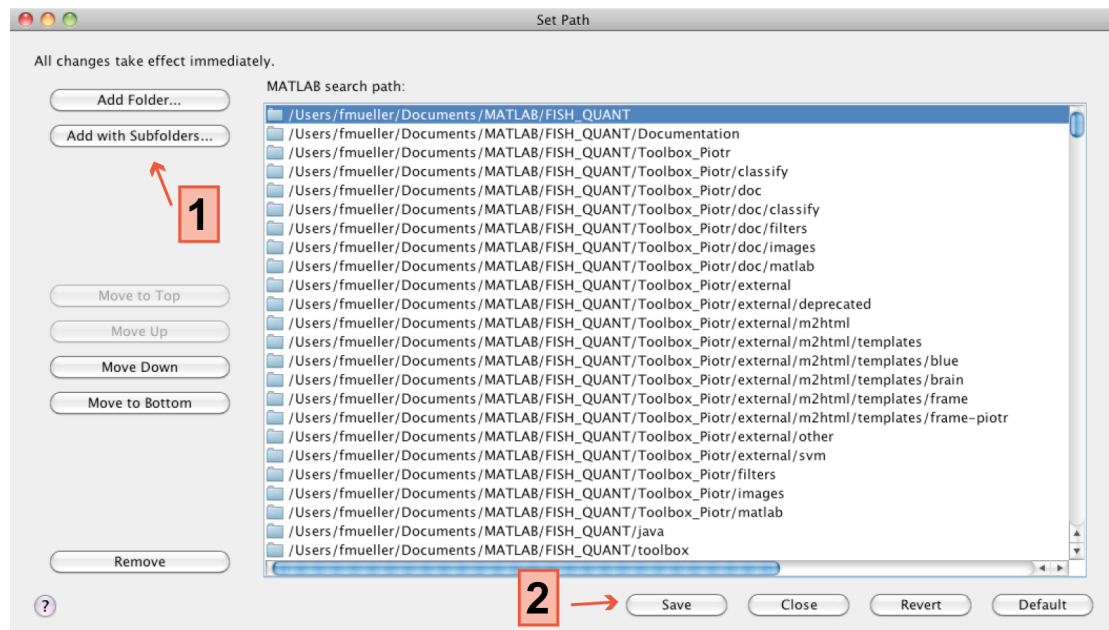
1.1. Install FISH-quant for Matlab

Install FISH-quant on the computer and setup Matlab to use it.

1. Download FISH-quant. Source code is provided on bitbucket https://bitbucket.org/muellerflorian/fish_quant
2. To download click on the 'Downloads' symbol on the left side  and select 'Down repository'. This will download a zip archive with all the source code.
3. Find user folder of Matlab. You can find this folder by opening Matlab and typing `userpath` in the command window.



4. Create a folder called 'FISH_quant' in the Matlab user folder. Extract the content of the downloaded archive into this folder.
5. Update path definition in Matlab. This step is very **important** since it will tell Matlab where to find the source code of FISH-quant.
 - a. In the Matlab menu select *File > Set Path*
 - b. This will open a dialog box. In this box select **Add with subfolders ...**
 - c. This will open another dialog. Here select the folder of FISH_QUANT from step 3, e.g. /Users/fmueller/Documents/Matlab. Click *OK*. The respective folders will be added (see screenshot below).
 - d. To save this settings press *Save*. Depending on the settings of the installation of Matlab this might results in a warning saying that the changes to path cannot be saved. Matlab proposes to save the path-definition file *pathdef.m* to another location. Click *Yes*. Select a directory of choice, e.g. the Matlab work directory of the user.



6. FISH-quant can then be started by typing FQ in the command window and pressing enter.

1.2. Typical files produced in FISH-quant - overview of example data set

The **example data** is again provided as a zip archive. It contains the typical files and folders that are generated during the analysis with FISH-QUANT. In this folder you can find the file **__FISH_QUANT_Test_data.pdf** containing a detailed description of all provided data.

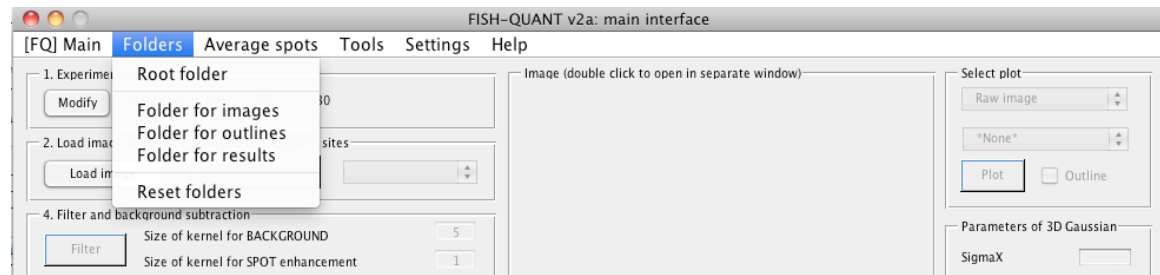
The results of the different analysis steps performed FISH-quant can be saved in different sub-folders. While these sub-folders are optional, we recommend using them since they help to keep the data organized. Several different folders can be used:

- **Images.** Here the FISH and DAPI images are stored. You can also save the filtered FISH images that would (FISH and DAPI), outlines (to define where the cells are in the images), and results (mature and nascent mRNA counting) are used.
- **Outlines.** Contains text files that contain the outlines of the individual cells in the images and well as the position of transcription sites and (optional) nuclei in these cells. The files also contain general information about the experimental settings and the images they are defined for. We provide manually defined outlines in the folder *outlines_manual* and automatically generated outlines files in the folder *outlines_automatic*.
- **Results.** Separate folders for the results of the mature mRNA detection and the transcription site quantification. The results file for the mature mRNA detection contain all the outlines, together with detailed information about each detected spots. Further they also contain the name of the images they were defined for. Different subfolders could also be generated to test different settings. Analysis is based on manually defined outlines.

1.3. Define a folder-structure for your project

As described before you can use multiple folders for the results generated in FISH-quant. The files describing outlines and results of mRNA detection contain information about the images they are defined for. HOWEVER, they do NOT contain information about the folders where these images can be found. These folders have to be manually defined in FISH-quant. This allows moving the files to other locations without having wrong references to folders in the files.

In FISH-quant you can define the different folders from the menu *Folders*. First you can define a root folder which will serve a starting folder for all the other folder explained above. For each folder a dialog box will open that will allow specifying the different folders.



Here you can either specify the folders of the example data or define the folders for the new project you want to work on. For the latter, we recommend generating the folders suggested above. This can be done in any file manager, e.g. Explorer under Windows or Finder under Mac.

1.4. Inspect example data

After defining the folders for the example data, you can play around with these files to get a first feeling for how FISH-quant works.

First, you have to define the folders.

1. Set root folder.
2. Set folder for images, outlines, and results (results_mature).

Things to do

- Open one of the outlines files and check out the outline-designer. For this go to *[FQ] Main > Load > Cell outline*. This will then open a dialog box and show the content of the outline folder. Select one of the files. FISH-quant will then open the associated image file and display the outline file. To change this outline, e.g. add or delete a cell, click on 'Define outlines'. This will open a new interface where the outlines can be changed.
- Open one of the mRNA detection files. For this go to *[FQ] Main > Load > Results of spot detection*. Here the folder with the results will be shown. Again select a file. FISH-quant will display the image, the outlines of the cells, and also all the detected spots. Now also the histograms in the lower part are populated. Here the fitted parameters of the 3D Gaussian fit are displayed for the currently selected cell (another cell can be selected in the pull-down menu next to the 'Define outlines' button). Changing the allowed range for the parameters will remove more or less spots. The detected spots can also be inspected in more detail in the spot inspector that can be launched by pressing on the 'Visualize spots' button.

The displayed results are also only for the currently selected cell. If multiple cells are defined for this image, they can be selected in the respective pull-down menu.

The screenshot displays the FISH-QUANT v2b main interface. The menu bar at the top includes [FQ] Main, Folders, Average spots, Tools, Settings, and Help. The interface is divided into several functional areas:

- 1. Experimental parameters:** Contains a 'Modify' button and fields for Sigma (x, y, z) with values 102 and 413.
- 2. Load image and define cells and transcription sites:** Includes 'Load image' and 'Define outlines' buttons, with a dropdown menu currently set to 'cell_1'.
- 4. Filtering, background estimation and SNR enhancement:** Features a 'Filter' button and input fields for Kernel BGD (pixels) and Kernel SNR (pixels), both set to 5.
- 5. Pre-detection:** Includes a 'Detect' button.
- 6. Fit spots:** Includes a 'Fit' button, a 'Restrict width' checkbox, and a 'Parallel computing' checkbox.
- 7. Thresholding:** Includes a 'Sigma - XY' dropdown, 'Min' and 'Max' sliders (values 127.6522 and 352.7919), an 'Apply threshold' button, and a 'Lock values' checkbox.
- Results of spot detection:** Displays two histograms: 'Selected # of spots: 72' and 'Total # of spots: 73'. Below these are 'Visualize spots' and 'Spot inspector' buttons.
- Image window:** Shows a microscopy image with detected spots (green circles) and cell outlines (yellow and blue). A legend indicates 'Fitted Spots' and 'Selected Spots'. A 'Select plot' dropdown is set to 'Raw image', and a 'Plot' button is available.
- Parameters of 3D Gaussian:** Includes input fields for SigmaX (172), SigmaY (172), SigmaZ (563), Amplitude (112), and Background (267), along with a 'Thresholded s...' button.
- Results of spot detection (double click to open in separate window):** Displays three heatmaps: 'Microprojection XY', 'Microprojection XZ', and 'RESID: Microproj XY'.
- Status bar:** Shows the progress of the detection process: 'PRE-DETECTION: FINISHED', '== STARTING matlabpool for parallel computing ... please wait ...', '... STARTED', 'Pre-detection in progress ... start up of GUI might take some time ...', 'Pre-detection: FINISHED', 'FITTING: STARTED ...', and 'Fitting: ... FINISHED'.

Annotations on the left side of the image point to specific features:

- Menu:** Points to the [FQ] Main menu item.
- Define outlines:** Points to the 'Define outlines' button.
- Select different cell:** Points to the 'cell_1' dropdown menu.
- Change thresholds for fitting parameters:** Points to the 'Sigma - XY' dropdown and the 'Apply threshold' button.
- Inspect detected spots:** Points to the 'Spot inspector' button.

Annotations on the right side of the image provide additional instructions:

- Image with FISH, outlines, detected spots:** Points to the main image window.
- Double-click on image to display it in separate window:** Points to the image window.
- Display elements can be selected in pull-down menu next to image:** Points to the 'Select plot' dropdown menu.

2. Defining outline files manually

One of the central text files in FISH-quant are the **outline files**. These files specify for each FISH image where the cells and their associated transcription sites and (optional) nucleus are in the images. The file also contains a reference to the FISH image it is associated to. FISH-quant allows loading (and associating) additional image files to an outline file

- i. DAPI image: can be used to automatically detect the nucleus
- ii. Image with an independent label of the transcription sites. Can be used to automatically detect transcription sites (explained below).

In the example data, three outlines files have been defined for the three images. They can be found in the outlines folder.

Defining the nuclei is not necessary for the analysis in FISH-quant. It can be useful to for the automated detection of transcription sites and determine the number of transcripts in the cytoplasm vs. the nucleus.

Usually outlines for all cells are defined before proceeding with the analysis.

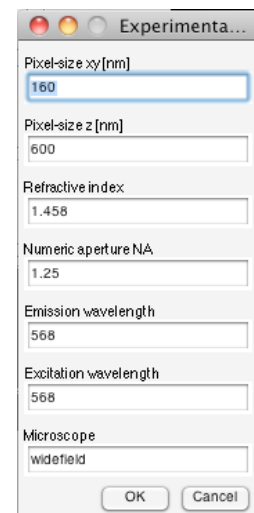
Example: the sub-folder outlines contains the outline files for three example images.

2.1. Define outlines

1. Define folders (at least root folder, and folders for images and outlines).
2. Define experimental parameters. Press 'Modify' to change a number of basic parameters. This has to be done **ONLY** once in an analysis session. FISH-quant will remember them for the subsequently opened images.

The plot on the right shows the parameters for the example data.

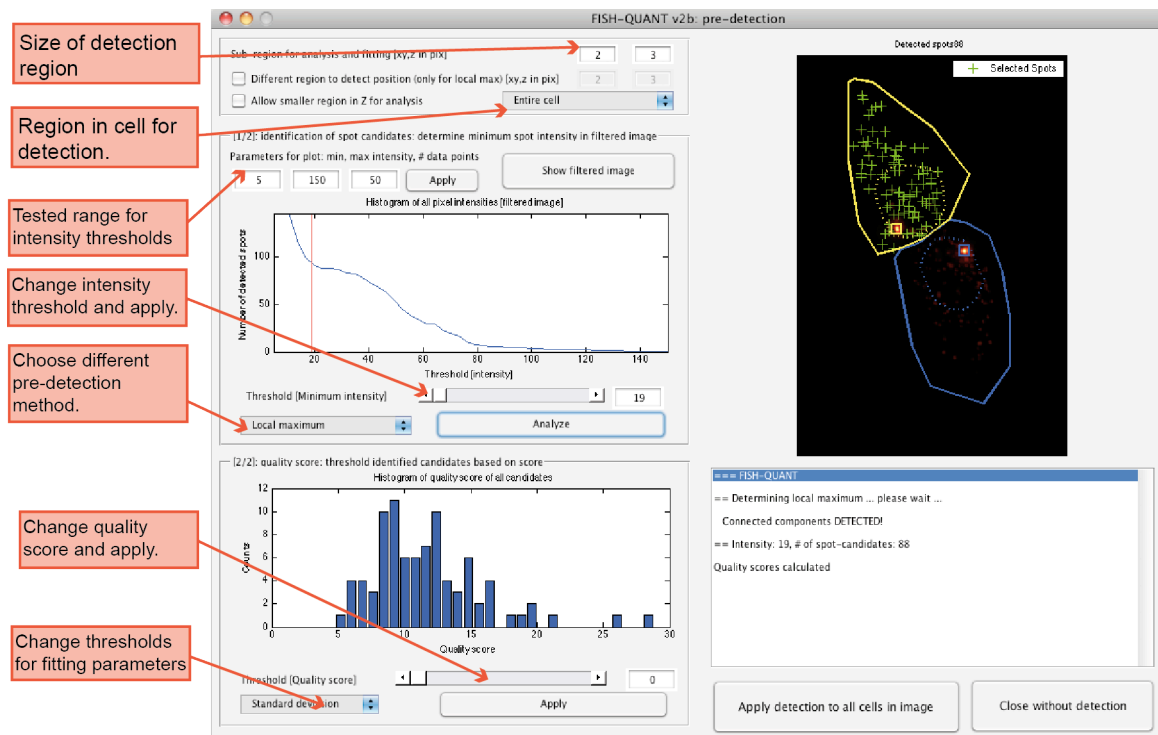
TIP: the default parameters can be changed. In the folder containing you can find a text file called "FISH-QUANT_default_par.txt" where you can change the values for each parameter. Change the value after the equal sign and don't add any spaces.



3. Open image. This can be done either from the menu *[FQ] Main > Load > Image [Raw]* or by pressing on the button 'Load image'. Click OK in the dialog that opens to confirm that the analysis results will be deleted.

Example: load image BGH_03_CY3.tif. This is stacked .tif file of mRNA FISH with two cells.

4. Press button 'Define outlines' to open separate user interface that allows defining outlines of cell and transcription sites (see below with a description of the basic controls). With these controls you can change the contrast, zoom, pan, and also switch from a view of the maximum intensity projection to a view of the individual z-slices.



5. Then you can define outlines. A typical workflow encompasses:

- Load DAPI image from the menu: *Load images > Load DAPI*. Image will be shown as an overlay with the FISH image. Either both or only one of the images can be selected (bottom left corner). Contrast of either image can be changed (controls under the plot). Contrasts can **NOT** be changed while you draw a region. Last, image can be inspected as a maximum intensity projection in xy or by navigating the different z-slices (Navigate 3D stack).
- Then you can define the outlines of the different cells. You can also choose the geometry of regions you want to use: polygons, ellipse, or rectangle. To draw a new cell press 'New cell', move over the image and wait until the cursor turns into a cross; draw the region; to **finish double-click** drawing the region (for polygon on the first point). Region will then be shown with a solid green line. You can then modify this region, or delete it if you are not satisfied.
- For each cell you can then define transcription sites and (optional) a nucleus. Draw the corresponding outlines and FISH-quant will automatically assign them to the corresponding cell. IMPORTANT: they have to be ENTIRELY within the cell.

Example:

Load image BGH_DAPI_03.tif. Change contrast of each image to optimize display. You can also toggle the display of the either image.

Then use the polygon tool to outline the two cells and their nuclei, and the rectangle selection to outline the two transcription sites.

- Press *Finished* to return to main interface. Save outline from the menu [*FQ Main*] > *Save* > *Outline of cells*.

To define outlines in a lot of different cells two different methods can be used

- Open each of the images in the main interface and perform the steps describe above.

- ii. Open the *Outline definer* from the tools menu in the main interface of FISH-quant. This will open the same interface as described above. However, now you can load the images from the menu *Load images > FISH*. Then you can define the outlines as describe and save them with *[FQ] outline > Save outline*. You can also add a comment. Alternatively, you can use the button 'Quick-save'. Here a default name will be proposed and no comment can be added.
- iii. Load a new image. Important: each time you load a new FISH image, the defined outlines will be deleted.

Important: when working with this option, define first the experimental parameters and the folder structure in the main interface. These settings will then be used in the *Outline definer*.

2.2. [Optional] Automatic detection of nuclei

Nuclei can be defined manually as described before or automatically based on a nuclear stain such as DAPI. Here FISH-quant proposes a threshold value (between 0 and 100). Press 'Detect nucleus' to perform detection with a given value. Nuclei will then be automatically. Depending on the signal this might not work for all cells, and the outlines of these cells have to be defined manually.

Example: for BGH_03_CY3.tif a value of 2.2 gives a good detection. Smaller values are not stringent enough and the two nuclei will not be separated. Larger values will be too stringent and the detected nuclei will be too small.

2.3. [Optional] Automatic detection of transcription sites

Transcription sites (TS) can be defined either manually or detected automatically. FISH-quant provides two different methods. In both approaches the detection is based on a minimum intensity threshold a TS has to have. In the first approach the FISH image is used directly, in the second approach an image with an independent label for the TS is used, i.e. LacI. For more details see the supplementary material (Note 4.4) of the Nature methods paper.

1. Load outline file or define cells in image. TS are only defined for the cells that are outlined, without cell no TS will be identified.
2. Optional: define parameters for TS detection from the menu *Options > TxSite detection*. The default parameters work often fine. Consult the FISH-quant user manual for a detailed description.
3. Define intensity threshold that separates transcription site from mature mRNA. This can be either done on the FISH image (select *FISH_image* in the pull-down menu) or on an image with a label for the TS (Select *TS_label*). This image can be loaded from the menu *Load images > Load TS-label*.

You can get a feeling for the values by selecting the *Cursor* tool in the lower part of the interface. When clicking on different parts of the image, a window will display the position of the cursor as well as the intensity of the FISH image and the second displayed image (either TS label or DAPI). Set the threshold value in the field next to *Threshold [TS]*.

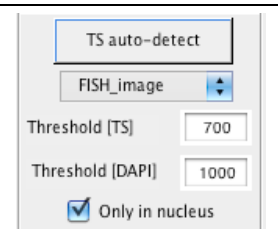
Press *TS auto-detect* to perform the detection. TS will only be detected in the defined cells. If the threshold is too low, then too many TS are detected. By default the detection stops when more than 100 TS are detected. Also the maximum number of TS per cells can be defined (by default 4). This can be changed in the options.

4. [Optional] Restrict detection to nucleus. To avoid over-detection, the detection can be restricted to the nucleus. Here two different methods can be defined.

- If nuclei are defined, then the option 'Only in nucleus' will restrict the TS to these defined nucleus. For cells where no nucleus is defined, then the entire cell will be considered.
- Define intensity threshold for DAPI signal. Rather than defining the nucleus, an additional intensity threshold for the DAPI signal can be set. This usually works well since the DAPI signal is substantially brighter in the nucleus. Simply define this value in the field next to *Threshold [DAPI]*.

Example:

For BGH_CY3_03.tif a intensity threshold of 700 for the TS, and 1000 for the DAPI gave good results.



TS auto-detect

FISH_image

Threshold [TS] 700

Threshold [DAPI] 1000

☒ Only in nucleus

The **TS detection settings** can also be saved and can be used in the batch mode for an automated TS detection. Menu *[FQ] Outline > Save settings for TS detection*.

3. Defining outline files automatically

In FQ v3, we introduced automated cell segmentation. For details refer to the separate user manual *FQ_cell_segmentation.pdf*. In short, we implemented a two-step process. First, the 3D stacks are projected into 2D images. Second, cell segmentation is performed on these 2D images and FQ outlines are generated from these segmentation results.

3.1. 2D Projections

This is done from a dedicated GUI that you can start with the command *FQ_seg* in Matlab. We used the following settings on the images.

The screenshot shows a MATLAB GUI titled "Make Z projection for segmentation". It is organized into several sections. The top section, "Slice selection based on global focus measurement", includes a slider for "% of kept best slices" set to 50 and a dropdown for "Focus operator" set to HELM. Below this is the "Parameters for Z-projection" section, which has two sub-sections. The first, "Standard projection", is unchecked and has a "maximum" dropdown. The second, "Local focus selection & projection", is checked and includes a dropdown for "Operator" set to HELM, a "Windows size" set to 7, a "Type" dropdown set to median, and a "# of slices" set to 3. The "Perform Z-projection" section contains two buttons: "Define images" and "Define folder". Below these are several options: "Search files recursively" (unchecked), "Image extension" set to .tif, "Ignore files containing" set to _filtered, "Save projections" (checked) with a "Name prefix" set to ZPROJ_, "Same folder" (unchecked) and "Sub-folder" (checked) with a ZPROJ field, and "Generate folder" (checked) with "images" and ">" fields. A large "Make projections" button is centered below these options. At the bottom, there are checkboxes for "Show plots" and "Store results for inspection" (both unchecked) and a "Show results" button.

- We selected the folder with the images inside (with the 'Define folder').
- This folder might already contain filtered images from an earlier processing run. To exclude these, we specified that files containing 'filtered' in their file-name should not be processed.
- The saving options might be a little confusing at the beginning, but help to get the files organized. To save the results, the program will replace in the entire folder name 'images' by 'outlines_automatic'. Within this 'outlines_automatic' folder, the program will create subfolder called ZPROJ where the 2D projections will be saved. So you essentially have now in your analysis folder a folder called *images* with the images, and *outlines_automatic* with a subfolder *ZPROJ* with the projections inside.
- To process, we clicked on 'Make projections'

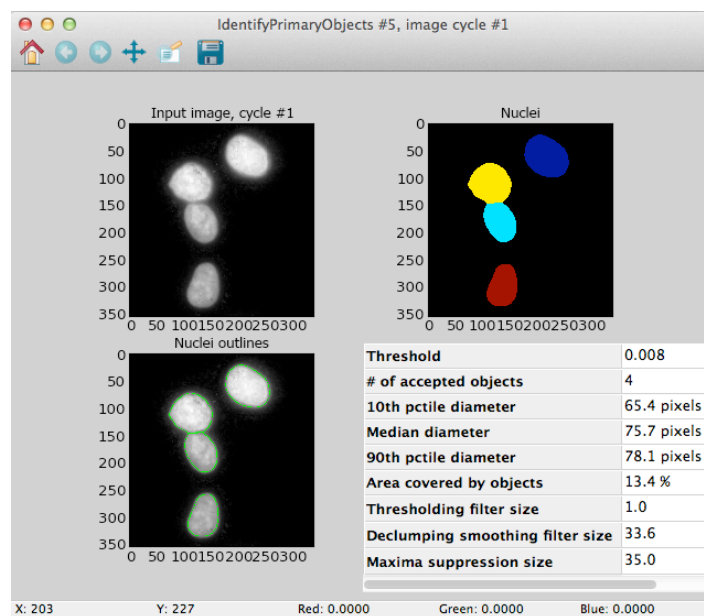
3.2. Cell segmentation

We provide instructions for how to use two different segmentation tools: *CellProfiler* and *CellCognition*. Please consult the help PDF for more details.

For this tutorial, we will use CellProfiler. The analysis is based on two steps (i) detection of the nuclei in the DAPI image (ii) the nuclei serve as seeds to find the cells in the smFISH image.

1. Open CellProfiler.
2. Import the analysis pipeline 'CP_pipeline.cppipe' that you can find in the folder outlines_automatic.
3. Drag and drop all the images (2D projections) into the dedicated window called 'file list'
4. Go the **Metadata** and click on the button Update in the window that opens in the right part of the interface.
5. Go to 'NamesAndTypes' and click on update. This will associate the CY3 and DAPI images.
6. You can then analyze the images in two main modes
 - a. Click on 'Analyze images' to analyze all specified images. To speed up the analyze close the 'eyes' next to each of the processing steps listed in the analysis module
 - b. Or, you can go into the test model. The latter allows inspecting the results of the different steps (most importantly the segmentation steps. The currently selected step is highlighted by a little slider on the left of the different analysis steps. When pressing 'Step' the highlighted step is performed and the cursor moved to the next step. To perform again the same step, simply move up the cursor.

The example below shows the output of the 'IdentifyPrimaryObjects' step, which finds the nuclei.



The results of the segmentation are saved in the same folder as the 2D projections. Two types of files are saved (i) files containing 'MASK'. These contain the actual segmentation masks, each identified object (cell or nucleus) has a unique gray value (ii) files containing 'CONTOUR'. These files show an overlay of the 2D projections and the identified objects. They can be used in a quick inspection to see how well the segmentation worked.

3.3. Automated outline generation

You now have to go back to the *FQ_seg* tool to generate the FQ outlines from the segmentation masks. The following steps are necessary

1. Define the experimental parameters
2. Define settings. See screenshot below. What I changed from default
 - a. Identifier for DAPI and smFISH and FISH
 - b. **Where to save the files:** Replace in the original folder 'ZPROJ' by ". This is NOT a typo. It simply means that ZPROJ will be replaced by an empty string. So you save the data in the folder hierarchy above, i.e. the 'outlines_automatic' folder.

Cell profiler: generate FQ outlines from segmentation

Define experimental parameters

Identifier for DAPI image:

Identifier for FISH image:

Identifier for nuclei segmentation:

Identifier for cell segmentation:

Part of file-name (projection) to remove:

☐ Generate outline files for 2nd color

☐ Don't generate for 1st color

☐ Same folder ☐ Sub-folder

☒ Generate folder

☐ Search files recursively

4. Mature mRNA counting

After the outlines are defined for all cells, the detection of mature mRNA can be performed. This analysis is one of the central parts of FISH-quant and it is necessary to understand the different parameters to optimize the detection. We therefore explain the different involved steps in detail below.

The mRNA counting will be done for the cells defined in the outline files. In the example data, the mature mRNA detection has been performed for the three outlines files and can be found in the folder *results_mature*. Here files summarizing the settings, and a results file for each of the outline files is saved.

The **overall workflow** can be split in two parts as also described below:

1. Define the detection settings for one cell (preferably with a lot of mRNA) and save the settings. This can be done on several cells to check that the settings are the same for each cell. If this is the case these settings can be used for a batch analysis.
2. The settings can then be used in a batch-processing module and applied to all cells. The final thresholding can then be done based on all detected spots in all images.

The mRNA detection itself involves **two major steps**

1. Pre-detection of spot candidates (mRNA) on filtered image.
2. Fit spot candidates with 3D Gaussian. This will be done on the raw image. Provides the position of each spot with high precision. Additional thresholding can be performed to separate true spots from noise.

4.1. Mature mRNA detection on individual images

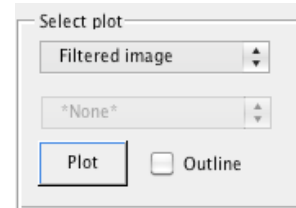
1. Continue from steps before or open FISH-quant, define folders, and open one already defined outline. Set the results folder as well (we usually recommend a separate folder for the results of the mature mRNA detection).

2. Filter image

We use a **two-step filtering**. First, the image is blurred with a Gaussian filter with a large kernel. This image is a good approximation of the background and is subtracted from the original image. The resulting image is then filtered with a small Gaussian kernel to increase the signal-to-noise ratio (SNR). The size of this filter is expressed in a multiplicative factor of the pixel-size. It is also possible to defined different Kernel size in XY and Z. Usually the same size is used for XY and Z. However, we found that using a smaller Kernel in Z for the background filtering can yield better results if only a few z-slices were acquired. The important thing here is that the resulting image has little background and bright, localized spots. By default the size for the large kernel (for the background) is set to 5 (for XY and Z), the size for the small kernel (for the increase of the SNR) to 1 (for XY and Z). Smaller values for this second Kernel could potentially give better results for images with spots that are very close together. Here you have to play around a little bit and see what settings work best.

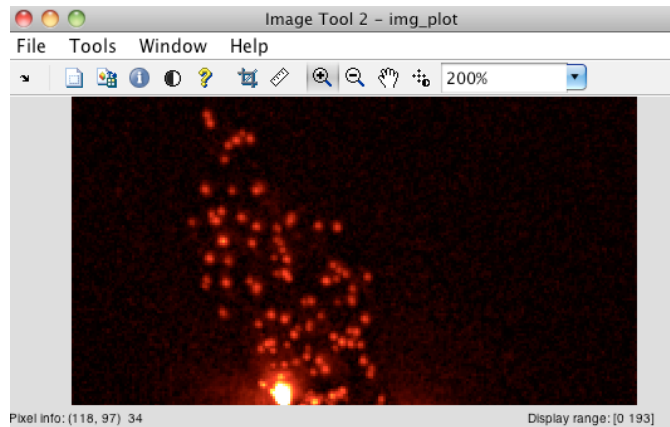
The filtered image can also be **saved** from the menu [FQ] Main > Save > Filtered image. It will be save in the defined folder for images. If you save the outline file now, a reference to this filtered image will be saved and the filtering step has not be done again. If the analysis is performed in the batch-mode (below) filtered images are generated and new outlines files saved that contain a reference to these images.

Tip: When you double-click on the image (**after** selecting filtered image, and disabling 'Outline') you can open the image in the *Image Tool* of Matlab. This allows inspecting the image by zooming, changing the contract,



See Matlab help for 'Interactive Exploration with the Image Tool' for more details.

When moving with the cursor over the image, Matlab displays the location and intensity of the pixel over which the cursor is currently located in the lower left corner (Pixel info).



Example: the default Kernel size works well for the test data. Change **these values to explore their impact on the filtered image. They might need to be adapted for your own data.**

3. Perform pre-detection

The pre-detection will be performed for the currently selected cell in a separate interface. These settings will then be applied to all cells in the image.

Before opening this interface, FISH-quant will test a range of threshold values and determine the number of detected spots for each tested value (explained in more detail below). This range has, however, to be adapted to your image data. The crucial parameter to set is the maximum intensity value that will be tested. To get this value we usually inspect the filtered image with the 'Image Tool' of Matlab (see above). By moving over the individual spots, you can get a first feeling for their brightness and determine what range the brightest ones have.

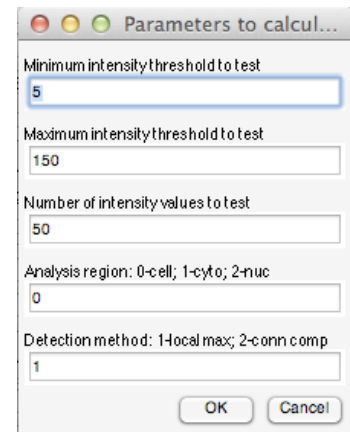
You can start the pre-detection by pressing 'Detect'. This will open a separate interface where the different parameters can be defined, before the interface is shown, a user dialog is displayed that allows specifying this range (below).

Additionally, the detection can be performed either for the entire cell, or only the cytoplasm or the nucleus. This requires that the outline of the nucleus is defined. For cells where no nucleus is defined, the detection will be performed in the entire cell, independently of the selection.

By default FISH-quant tests 50 values between 5 and 150. Here the maximum value can then be adjusted (to around 1.5-2 times the intensity of the brightest spot).

The next entry allows specifying in **which part of the cell** the detection will be performed: 0 - entire cell; 1 - only cytoplasm; 2 - only the nucleus. This can also be changed afterwards in the GUI.

The last entry allows specifying the **pre-detection method**: 1 for local maximum [default] or 2 for connected components.



Parameters to calculate...

Minimum intensity threshold to test: 5

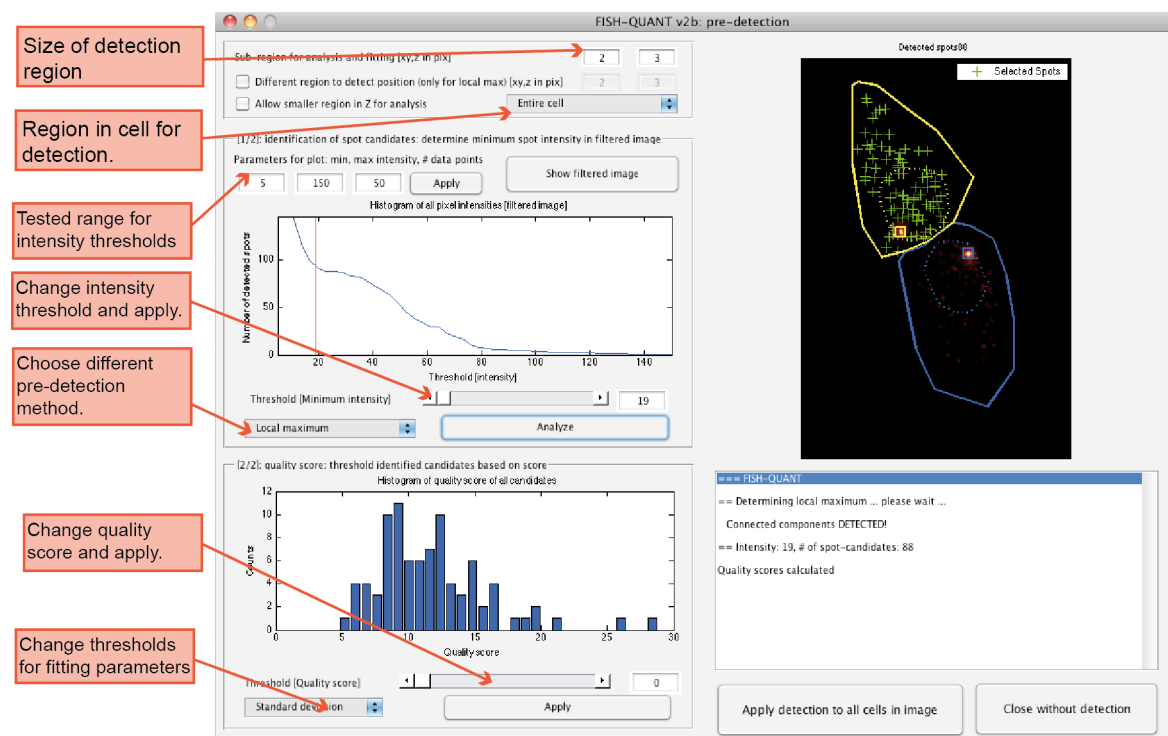
Maximum intensity threshold to test: 150

Number of intensity values to test: 50

Analysis region: 0-cell; 1-cyto; 2-nuc: 0

Detection method: 1-local max; 2-conn comp: 1

OK Cancel



FISH-QUANT v2b: pre-detection

Sub-region for analysis and fitting (x,y,z in pix): 2 3

☐ Different region to detect position (only for local max) (x,y,z in pix): 2 3

☐ Allow smaller region in Z for analysis: Entire cell

[1/2]: Identification of spot candidates: determine minimum spot intensity in filtered image

Parameters for plot: min, max intensity, # data points: 5 150 50 Apply Show filtered image

Histogram of all pixel intensities (filtered image)

Number of detected spots: 100

Threshold (Intensity): 19

Local maximum Analyze

[2/2]: quality score: threshold identified candidates based on score

Histogram of quality score of all candidates

Count: 12

Quality score: 0 5 10 15 20 25 30

Threshold (Quality score): 0

Standard deviation Apply

Detected spots: 88

Selected Spots

== FISH-QUANT ==

== Determining local maximum ... please wait ...

Connected components DETECTED!

== Intensity: 19, # of spot-candidates: 88

Quality scores calculated

Apply detection to all cells in image Close without detection

Size of detection region

Region in cell for detection.

Tested range for intensity thresholds

Change intensity threshold and apply.

Choose different pre-detection method.

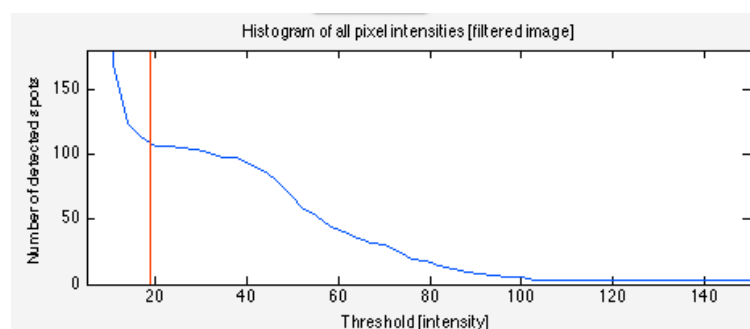
Change quality score and apply.

Change thresholds for fitting parameters

FISH-quant will then perform the pre-initialization (see also workspace for progress) and show the pre-detection interface. Depending on the size of the image this can take a few seconds. A number of important settings are defined in this interface:

- **Size of the detection region** around each detected spots. A region of this size will then be cropped and used to fit the 3D Gaussian. Spots are by default NOT considered if such a region can't be extracted (because they are too close to the edge). This region is calculated based on the theoretical size of the PSF. Usually these values are ok but they can be adjusted in case they are too small (not enough of the signal is considered) or too large (too much background is considered and/or a lot of spots are rejected because they are too close to the border of the image).
- **Intensity threshold for the pre-detection.** This is the minimum value a spot must have in the filtered image to be considered. FISH-quant shows a plot that help to select this threshold. The plot in the panel [1/2] (see below) shows the number of detected spots as a function of the used intensity threshold. For a low intensity threshold too many spots are detected, for a high threshold only a few spots are detected. For the range that

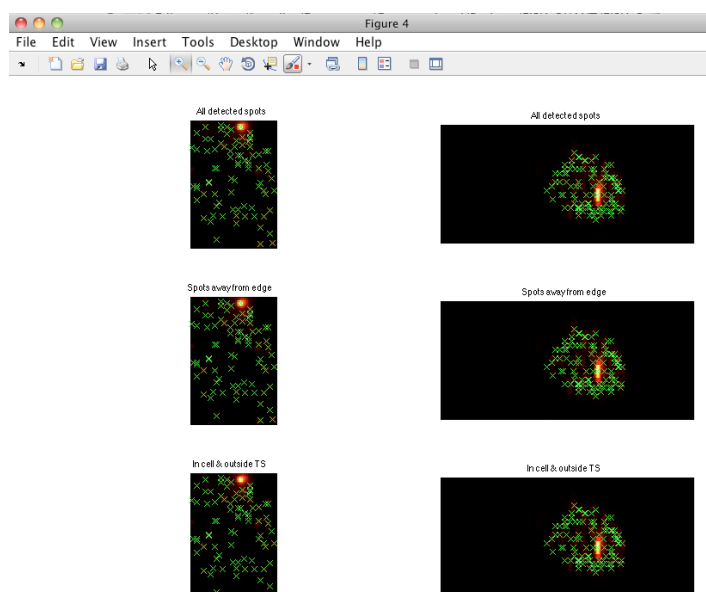
corresponds to the individual spots a plateau is usually visible. Place the threshold towards the left side of this plateau (A value of 19 for the plot below).



By pressing 'Analyze' the selected threshold value will be applied. FISH-quant will then display an image that shows the results of this detection for the selected cell.

This plot displays the detected spots in maximum intensity projections in XY (first column) and XZ (second column).

The first row shows all spots. The second row shows spots that are away from the edge, i.e. the detection region can be cropped around them. The third row shows spots that are away from transcription sites.



You can zoom-in the different plots. The three plots in a column are connected, so if you change the zoom in one of them, the others will be updated. This guarantees that always the same field of view is displayed. These plots can already give an indication if a good threshold was chosen. FISH-quant provides two different pre-detection methods, they usually provide similar results (see manual for more detail).

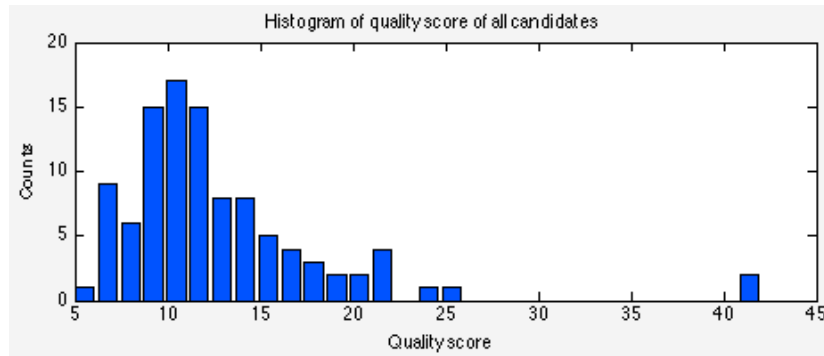
Example: a value of 19 gives good results. A smaller value leads to over-detection, a much larger value to under-detection.

- Optionally an additional quality score can be determined. Here two different methods are implemented (i) standard deviation of all pixel intensities in the cropped area around each spot (ii) 3D curvature. The latter is especially suited for brighter FISH data. Either metric gives larger values for bright spots which are usually individual mRNA. An additional threshold can then be defined for this quality score.

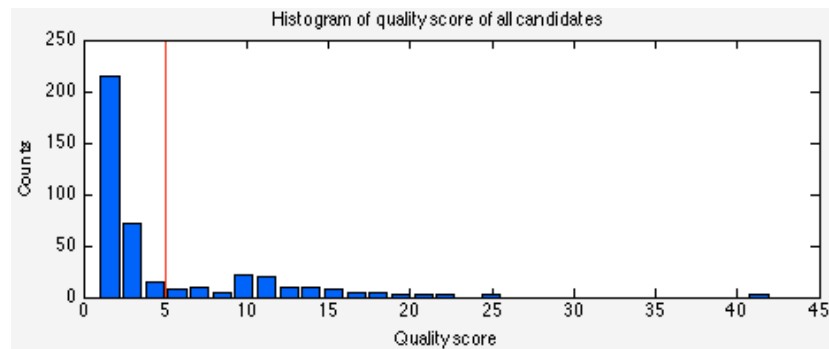
Depending on the quality of the image and how well the pre-detection worked, this threshold might not be necessary. We usually recommend setting the pre-detection threshold in a way that false-positives are already avoided as much as possible. If necessary, the quality score can then be used as well. A final thresholding can then be performed based on the fits with the 3D Gaussian (below).

Example:

With a detection threshold of 19 mainly good spots are detected (around 100 mRNAs for Cell_2). The quality score histogram shows one broadened peak that starts at around 5 (for the standard deviation).



Decreasing the detection threshold to 5 leads to a massive over-detection (around 400 mRNAs). However, these false-positives have a smaller quality score. By setting the threshold for the quality score to 5, most of them are rejected (only 108 mRNAs pass this test).



Similar results can be achieved with the quality score based on the curvature. Here the difference in the histograms is even larger.

- To apply the setting to all cells in the image press '*Apply detection to all cells in the image*'. The interface will then close and the pre-detection be performed. Depending on the number of cell in the image, this can take a few minutes. A status update is provided in the Matlab command window.

4. Fit detected spots with 3D Gaussian.

The detected spots are then fit with a 3D Gaussian. This is performed on the raw image since filtering affects the intensity in a non-trivial way and also affects the localization precision.

On computers with multiple cores this analysis can be performed with **parallel computing** (requires the Matlab Parallel Computing toolbox). This option can be enabled by enabling the option '*Parallel computing*'. This takes a few moments, so please wait until the status update indicates that the activation is completed.

To start the fitting press '**Fit**'. Depending on the number of detected spots this can take a few minutes. Updates are provided in the Matlab command window. When the fit is then a number of windows will be populated in the lower part of the main interface.

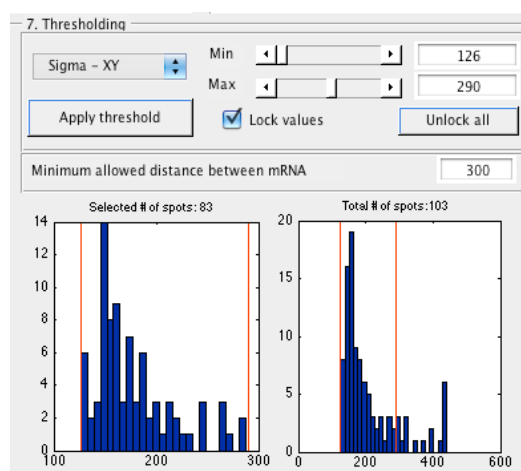
The panel with the label '**7. Thresholding**' allows to inspect the different (fitting) parameters that were estimated for each spot. They can be used to further exclude spots that are false-positives and are background. The relevant parameters here are the size of the spot in XY ('Sigma-XY'), the size in Z ('Sigma Z'), and the amplitude ('Amplitude'). Background spots

have usually large sigma's and small amplitudes. To exclude them, the user can define for each parameter upper and lower boundaries.

- Select the parameter you want to restrict. The two histograms in the lower part will show the distribution of this parameter. The right histogram of all spots, the left histograms only for the spots that passed the thresholding. These histograms are shown for the detected spots in the currently selected cell.
- Define the boundary for the upper and lower threshold. This can be done by either moving the slider or changing the values in the text field. The red vertical lines move accordingly and indicate the current value of the thresholds.
- Press 'Apply threshold' to make this selection active.
- Enable 'Lock values' to make this selection permanent. Now you could set additional thresholds with other parameters. Pressing 'Unlock all' will remove all defined thresholds

Example:

Thresholds for sigma-xy are set to a minimum of 126 and a maximum of 290. The threshold values are also locked.



Spots stemming from noise usually have a large sigma, and small amplitude. So they can be relatively easily identified. The distribution of the sigma's and the amplitude for the real spots looks like a skewed Gaussian distribution, while the values from noise are outliers to the left and the right of this distribution.

It is interesting to perform this analysis on individual cells but we found that the histograms become more informative when they are plotted for all detected spots in all cells. This can be done in the batch-processing interface (explained below). In this interface the same controls are available to define thresholds on the estimated parameters as described here – except that it is done for all detected spots in all cells at the same time.

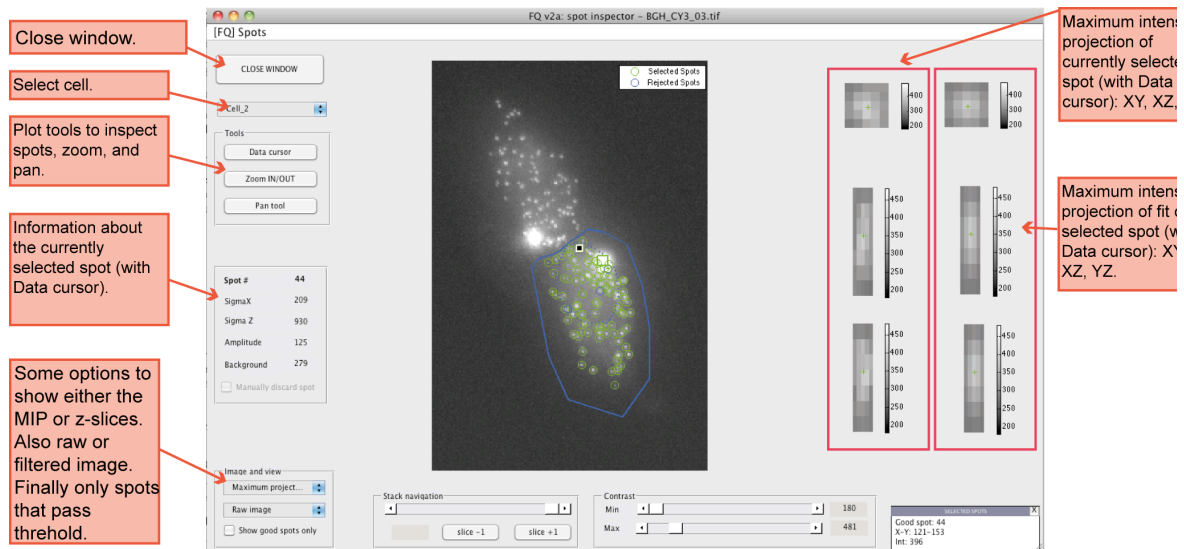
5. Inspect detected spots

The detected spots can then be inspected with the 'Spot inspector' by pressing on '*Visualize Spots*' in the lower part of the interface.

In this interface the detected spots in each cell can be displayed. Spots that pass the thresholding of the step before are shown in green, spots that fail in blue. There are various general controls to change the way the image are displayed that are identical to what can be done with the outline definier. Additionally, the user can display either the raw or the filtered image.

The tool 'Data cursor' can be used to inspect the individual spots. After selecting it you have to DOUBLE-CLICK on the spot you want to inspect. FISH-quant will then display the estimated parameters of this spot (left side of the interface). On the right side of the

interface FISH-quant will then display maximum intensity projection (MIP) along the three axis of the spot in the image (first column), and of the fit with the 3D Gaussian (right column).



6. Save results of spot detection

Results can be saved in a text file from the menu *[FQ] Main > Save > Detected spots*. Save for each spots all acquired information. Two options are available: either save all spots, or only the spots that pass the thresholding. We usually choose the former since we can then adjust the thresholding later if needed.

Example: the three files in the results folder ending with *_outline_spots_130305* are such files.

7. Save settings for batch processing

The final step is to save the specified detection setting for future reference and also to use them in the batch-processing module. This can be down from the menu *[FQ] Main > Save > Detection settings*. This will generate a text file containing the experimental parameters, and the settings for filtering, pre-detection, and fitting.

Example: see the file *_FQ_settings_MATURE.txt* in the *results_mature* sub-folder.

4.2. Mature mRNA detection in batch mode

If the detection settings are comparable for different cells, then the detection can be performed automatically for all outlines files. Images have NOT to be filtered the filtering is done also in the batch-processing. This means that the user has to provide outlines files for all images, determine the settings, and then the processing can be done automatically.

This is done in the batch-processing module of FISH-quant. This module is also used for the transcription site quantification, but here we will only focus on the relevant parts for the mature mRNA detection.

1. **Open batch processing module** from the menu of the main interface *[FQ] Main > Tools > Batch processing*. You can then also activate the parallel computing capability if this is possible on your system.
2. **Define folders** as in the main interface. If the folders are already defined in the main interface you can set the same folder for the batch-processing from the menu *Folders > Use Folders from FQ-main*.

3. **Load settings file** (Panel '1. Define file with settings') . Here you load the file that you save in the preceding section

Example: load the file `_FQ_settings_MATURE.txt` in the `results_mature` sub-folder.

4. **Load the outline files** to be processed (Panel '2. Outlines to be processed'). Here you can multiple files – but they have to be in the same folder.

Example: select the three outline files from the `outlines` sub-folder.

The screenshot shows the FISH-QUANT v2b: batch mode interface. Red boxes with arrows point to specific features, each with a text annotation:

- Load analysis settings:** Points to the 'Load' button in the '1. Define file with settings' panel.
- Define outline files that should be processed:** Points to the 'Add' button in the '2. Outline files to be processed' panel.
- Also select files here to be opened with Spot inspector:** Points to the 'Files listed' button in the '2. Outline files to be processed' panel.
- Start batch-processing:** Points to the 'PROCESS' button in the '3. Process data (detect mature mRNA)' panel.
- Controls to determine if filtering should be performed or not:** Points to the 'Use filtered images' checkbox in the '3. Process data' panel.
- After thresholding is applied, effect on different cells is displayed:** Points to the 'Intensity threshold' slider in the '4. Thresholding' panel.
- Threshold spots based on estimated parameters of fit:** Points to the 'Quantify TxSite' button in the '5.2 Quantification' panel.

5. The batch-processing will also perform the filtering if

- No filtered image is present
- The option 'Use filtered image' is disabled.

Usually we don't do the filtering for each image, but define only the outlines and the settings and FISH-quant will then do the filtering automatically. This saves a lot of time since the filtering can take a lot of time.

By enabling the option 'Save filtered image' the filtered image will be saved in the image subfolder. FISH-quant will then also generated new outline files with a reference to these images. When opening these files, FISH-quant will automatically open the filtered image as well.

To separate these new outline files from the old ones, we make a new folder to save these files. We then update the outline folder in the batch-processing to this folder. The new files will then be saved in the new folder. This will not affect the files that are already defined in step 4 above.

Example:

See subfolder `outlines_batch`. Contains the same outlines as the subfolder `outlines`, except that there is a reference to the filtered images. These outlines also have the suffix `_batch`.

The filtered image are indicated with the suffix `_filtered_batch`.

6. Run analysis by pressing 'Process'.

Depending on the number of files this can take between a few minutes to a few hours. Updates are provided in the interface and the Matlab command window.

When the analysis is done two histograms will show up. As for the analysis of individual cells they can be used to display the distribution of one of the parameters estimated for the individual mRNA spots. However, here the histograms are populated with the values of all

spots from all cells. Usually it's much easier to define good threshold values. Same considerations for the thresholding apply as mentioned above.

Each time a threshold is updated, it will be applied to all cells. The effect on each image is indicated in the status window where for each cell the number of mRNA spots before and after the thresholding is shown.

The detection can also be verified in the images: select the outline you want to look at in the list in panel 2. Then press on 'Show detected spots'. This will open the spot inspector window described above.

Example: we use the following min/max threshold for three parameters: sigma-xy: 117/240; sigma-z: 360/800; amplitude: 50/250.

7. Save results.

A number of different results files can be saved from the menu *[FQ] Batch > Save*. The results will be save in the specified results folder.

- i. *Settings [mature detection]*: this will save the analysis settings. Especially interesting if you change the detection thresholds.
- ii. **Results for each image**: this saves a results file for each outline file that containing all information about each detected spot in the cells defined in this outline. Two different options are available: either all detected spots or only the spots that passed the final thresholding. We usually use the former option since this way we can inspect all spots and potentially adjust later the thresholding settings. Such a text file is generated automatically for each outline file.

Example: the three files in the results folder ending with `__outline_spots_130305` are such files.

- iii. **Summary: mature mRNA**. Generates one text file containing a summary of the number of detected spots for each cell. The file lists the total number of detected spots, the number of spots that passed the threshold, and (if defined) the number of spots in the nucleus.

Example: see the file `__FQ_batch_summary_MATURE_130305.txt` in the *results_mature* subfolder.

5. Nascent mRNA counting

FISH-quant also allows quantifying the number of nascent transcripts at the transcription site (TS). FISH-quant compares the signal of the TS to the typical signal of mature mRNAs and infers the number of equivalent full-length transcripts at the TS, i.e. how many mature transcripts would give the same signal as observed at the TS.

This analysis requires that the signal of the individual mRNA is well defined. We therefore use the results of all detected spots in all cells. This usually gives several thousand individual mRNA's that are considered in the analysis. With fewer spots not all methods might work reliably.

Several different methods are implemented. The first two methods consider the spatial extent of the transcription sites and are the method of choice.

- i. **Integrated intensity.** Here the TS is also fit with a 3D Gaussian and the integrated intensity under the curve is compared to the integrated intensity for the mature mRNAs. **This approach should work for most transcription sites.** It starts to fail only for elongated, ellipsoidal sites. We recommend this approach for most applications and explained in this document.
- ii. **PSF superposition approach.** Here the image of the TS is reconstructed. In an iterative process the image of an individual mRNA are placed one after the other until the best description of the TS is obtained. This takes longer than the integrated intensity approach is most of the time not needed. For more information see the user manual of FISH-quant.

The two other methods ignore the spatial extent, and work well for relatively small TS that are not much larger than individual transcripts (see also the Nature methods paper describing FISH-quant for more details).

- iii. **Amplitude.** TS is fit with 3D Gaussian and the estimated amplitude is compared to the amplitude of the individual mRNAs.
- iv. **Maximum intensity.** The intensity of the brightest voxel of the TS is compared to the brightest voxel of the individual mRNAs.

The overall workflow is comparable to the mature mRNA quantification. First, the settings are defined for one example. Then all outline files are processed in the batch-processing module.

5.1. How to obtain the necessary files for the quantification

To perform the TS quantification information about the individual mRNAs is necessary. This is usually obtained in the batch-interface after the mature mRNA quantification is finished (as described above).

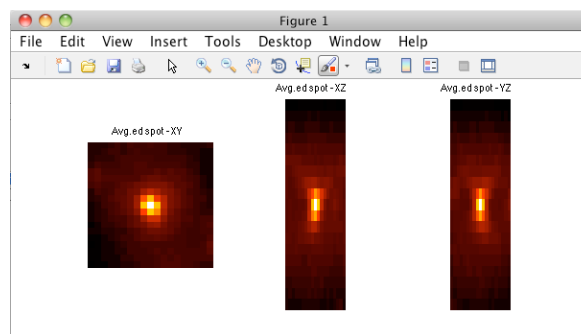
Alternatively the results can be loaded into the batch interface *[FQ] Batch > Load > Results of spot detection*. Here the results files for each image can be loaded into the batch-interface.

Example: load the three files in the subfolder *results_mature* ending with *__outline_spots_130305*. This will automatically also load the associated settings file.

For the quantification with the integrated intensity you only need the average image of the mRNA

This image is obtained by averaging the images of all individual mRNA molecules together. It will then be used to determine the integrated intensity and it will also be used in the PSF superposition approach.

To average all mRNA images together go the menu item *Average spots > Average all thresholded spots*. Then a dialog box will open with some parameters describing the averaging (for this example simply confirm. The most important parameter to the size of the cropping region in z. This is expressed in σ – too large values will lead to a rejection of a lot of spots because they are too close to the edge of the image. For more information consult the user manual.). Then all spots will be averaged. When the averaging is finished, a window will be displayed that shows the averaged image as a maximum intensity projection along the different axis (see below). This image is still quite noisy since only a few hundred spots are averaged together. The image quality will improve if more spots are averaged.



The image can then be saved as a TIF file from the menu *Average spots > Save as TIF > Normal sampling*.

Example: see image *_mRNA_AVG_ns.tif* in the subfolder *results_nascent*. This image was obtained by averaging the image of 6K individual mRNA molecules.

5.2. Transcription site quantification with integrated intensity

Here the settings for the transcription site quantification are defined. Start from the main interface with an outline file opened, e.g. the outline file for BGH_CY3_03.tif.

1. **Change results folder** in the main interface to a new folder containing the results of the TS quantification, e.g. *results_nascent*.
2. **Open transcription site module.**
From menu in main interface *Tools > TxSite quantification*.
3. Import **averaged image of mRNA** molecules.
Press button 'Image of mRNA'. Here you load the average image of the mRNAs.

Example: define image *_mRNA_AVG_ns.tif* in the subfolder *results_nascent*.

4. **Perform quantification.**
Then the TS quantification can be tested either on the currently selected TS (by pressing 'Selected site') or on all TS in the image (by pressing 'All sites').

By default the PSF superposition approach is not applied. It is computationally more intensive than the integrated intensity approach, and for most cases not necessary since the integrated intensity approach works fine. But it can be enabled with the respective toggle button.

The TS quantification also supports **parallel computing** on suitable systems. This option can also be enabled with the respective toggle button.

Example: quantify the TS in *Cell_1* (called *TxS_auto_2*) by first selecting it, then enabling 'Use also PSF superposition approach', and finally pressing 'Selected site'. Will give the following results (as displayed in the Matlab command window. Here the estimated number of nascent mRNA at the TS with the different methods are listed

```
==== SUMMARY OF TxSITE QUANTIFICATION ====

All REC: mean +/- stdev      : 46 +/- 2
Best 10% of REC: mean +/- stdev : 48 +/- 3

Number of mRNA at TS (ratio integrated INT): 42
Number of mRNA at TS (sum of pixel)       : 30
Number of mRNA at TS (ratio max. int. )   : 20
Number of mRNA at TS (ratio fitted AMP)   : 20

AVERAGED SIZE OF TRANSCRIPTION SITE
All mRNA : 594 +/- 24 nm
mRNA within max distance : 594 +/- 24 nm
```

5. Save settings.

Then the settings can be save from *[FQ] TxSite > Save settings*. This will generate a text file with the settings that can then be used in the batch processing-module.

Example: see file *_FQ_settings_NASCENT.txt* in the subfolder *results_nascent*.

There are a **number of other settings** that can be changed for the TS quantification. The most important one is the cropping size around each TS. FQ automatically crops a 3D regions around each the brightest voxel in each TS for the analysis. The size of this region can be change from the Menu. The cropping size is defined in +/- nm in XY and Z. This window should be large enough to contain even the large TS in your study but also small enough to not contain too much background. We recommend adjusting the size while looking at typically small and large TS. A discussion of the remainder of the parameters is beyond the scope of this document; please consult the user manual for more details.

5.3. Transcription site quantification in batch mode

1. Start interface.

The TS quantification settings can now be applied in the batch-processing module. Load this module as for the mature mRNA quantification. Set the correct folders. Load the mature mRNA settings (necessary for the experimental parameters).

Then the outline files that are processed have to be specified. Here the best is to work after the mature mRNA quantification is performed. This allows saving a summary file containing the results of the mature and nascent mRNA quantification. This means that the TS quantification is done either directly after the mRNA quantification is performed, or the mRNA quantification results files are loaded in the interface as described above.

Example:

Steps 1-2 have only to be performed if the batch interface does not show the results of the mature mRNA quantification.

- i. Set the relevant folder structure to the example data.
- ii. **Load the three files** in the subfolder *results_mature* ending with *_outline_spots_130305*. It is possible to select multiple files at once. **DONT** load the file summarizing all detected spots. This file doesn't contain any information about the cells and where the transcription sites are.

2. Load TS quantification settings

This is done with the button 'Load settings' in the panel '5. Quantify TxSites [nascent mRNA]'.

Example: load file *_FQ_settings_NASCENT.txt* in the subfolder *results_nascent*.

3. **Analyze settings:** press button 'Analyze settings'.
Will take again a few moments to perform the pre-analysis. Updates are provided in the command window of Matlab.
4. **Perform analysis:** press 'Quantify TxSite'
You can again choose if you want to use the PSF superposition approach or not and if you want to use the parallel computing option. There are a number of options that allow saving some more detailed results about the quantification (we won't use this here – see manual for more details. Will quantify all TS in all cells.
5. **Restrict size**
This affects the PSF superposition approach. The size of the TS that will be considered in the analysis can be restricted. All mRNA that are outside of this radius (measured from the center of mass of the TS) will not be considered as nascent transcripts. This helps to reduce the effect of mRNA's that are placed outside the TS to describe background or individual mRNA's that are within the measurement area.

Example: restrict size to 1000 nm.

6. **Save results files.**
Different results files can be saved from the menu *[FQ-Batch] > Save*.

Summary: nascent mRNA

Gives a detailed summary for each transcription site. Here the results of the different quantifications together with the estimated size are saved.

Example: see file *_FQ_batch_summary_NASCENT_130306.txt* in subfolder *results_nascent*.

Summary: nascent and mature mRNA

This option is available if the amount of mature mRNA has been quantified for each cell. Then a summary file for each cell containing the amount of mature and nascent mRNA can be generated. The user can decide which result of the nascent mRNA quantification will be used in this file. **This file is well suited as an input for subsequent kinetic modeling.**

Example: see file *_FQ_batch_summary_ALL_130306.txt*. The transcription site quantification is based on the integrated intensity.