FISH-quant v3

Matlab package to analyze and visualize data of mRNA FISH experiments

A depot of the software was issued at APP (http://www.app.asso.fr/en/) under the reference number IDDN.FR.001.090009.000.S.A.2013.000.10000

For more details see the **publication**: Mueller F, Senecal A, Tantale K, Marie-Nelly H, Ly N, Collin O, Basyuk E, Bertrand E, Darzacq X, Zimmer C. *FISH-quant: automatic counting of transcripts in 3D FISH image*. Nature Methods. 2013; 10(4): 277–278.

Florian Mueller, muellerf.research@gmail.com
Institut Pasteur Paris, Computational Imaging & Modeling Unit
25-28 rue du Docteur Roux, 75015 Paris, France

1. Installation

Requirements

FISH-quant was developed and tested in **Matlab R2014a** on Mac OS 10.9.5. Some function might **NOT** work in earlier versions (see section on Troubleshooting for known problems). Please contact us if you encounter any other problems. Several function of *the Piotr's Image & Video Toolbox for Matlab*[1] are used. The following Matlab toolboxes are needed in addition

- *Optimization Toolbox*
- *Statistics Toolbox*
- *Image Processing Toolbox*
- *Parallel Computing Toolbox* (***Optional***, speeds up processing time on computers with multiple cores)

Downloading and installing the Matlab code

1. Download code from Bitbucket in the downloads section (Select *Download repository*) https://bitbucket.org/muellerflorian/fish_quant/overview
2. The code is provided in a zip archive with a name like *muellerflorian-fish_quant-b4177b99dc53.zip*. The last part of the name will change over time reflecting new version. To avoid updating the path-definition each time you download a new version of FISH-quant (see next), we recommend copying the content of this archive to a folder called FISH_quant in the user folder of Matlab.
3. Under windows this folder is usually `C:\Users\user_name\Documents\MATLAB`, where `user_name` is the user name. This path can be found with the Matlab command `userpath`.
4. Create a folder FISH-quant, and copy the content of the downloaded archive in this folder. In the example above, you will have a folder `C:\Users\user_name\Documents\MATLAB\FISH_QUANT` with all the source code.
5. **Update Matlab path definition**. This can be done with a few simple steps in Matlab.
    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 1, e.g. `C:\Users\user_name\Documents\MATLAB\FISH_QUANT`. Click *OK*.
    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.

IMPORTANT 1 – FISH-quant on 32 bit operating systems

FISH-quant uses several routine from the Piotr's Matlab toolbox[1]. The distributed version has already all the necessary files compiled from 64 bit Windows/Linux version. For 32 bit operating systems you have to compile some files once before using FISH-quant. Open Matlab and type in the command window 'toolboxCompile' (without the apostrophe) and hit enter.

IMPORTANT 2 – older version of FISH-quant

If an older version of FISH-quant is already installed on the computer, either delete this folder, or move it to a different location.

Starting FISH-quant

Type **`FISH_QUANT`** or **`FQ`** in the command window.

Note: Don't use CLEAR ALL while working with FISH-quant. Some of the parameters are declared as global and will be deleted with this command. FISH-quant will not work afterwards.

Bio-Formats

FISH QUANT uses the function `bfopen` to read-in image data[2]. *Bio-Formats is a standalone Java library for reading and writing life sciences image file formats.* This allows reading in a wide range of different image formats. For a list of supported format consult the Bio-Format web site. FISH-quant currently supports only single color stacks.

[*OPTIONAL*] MIJ – Interface between Matlab and ImageJ

FISH-quant can display some of the obtained images in Fiji (ImageJ)[3] by utilizing MIJ[4]. To use this functionality a few simple initialization steps have to be performed once. A more detailed description can be found on the website http://fiji.sc/Miji

1. FIJI has to be installed properly according to the instructions on the website http://fiji.sc.

2. Add the scripts directory of Fiji to the to the Matlab path definition. On MAC this is not directly possible so use the command *addpath('/Applications/Fiji.app/scripts')* instead. You have then to SAVE the updated path definition. Go to the menu *File > Set Path* and select save.

If large images will be used the Java heap memory has to be adjusted. Increasing the size of the Java memory may correspondingly decrease the amount of space available for MATLAB arrays. There are two different ways to achieve this. Check the Matlab page for details.

- With more recent versions of Matlab, you can increase the memory by using File > Preferences > General > Java Heap Memory.
- It this is not possible a file called *java.opts* file can be either generated or changed. This file can be found $MATLABROOT/bin/$ARCH directory. $MATLABROOT is the installation directory of Matlab, /$ARCH is the architecture of the system, i.e. maci64 for a 64bit Mac OS. To increase the memory to 512 MB add the following line to the text file "-Xmx512m". If you don't have write access to this file you can save a local copy in the user path (see above for how to find the user path).

[*OPTIONAL*] 3D rendering with matVTK

It is possible to show the detected spots together with the image in a 3D rendered plot. FISH-quant uses matVTK[5] for this visualization. In order to use this functionality this software package has to be installed according to the instruction on http://www.cir.meduniwien.ac.at/matvtk/. The path with the Matlab client scripts has to be added to the Matlab path definition as described above.

2. Overview

Summary

FISH-quant is Matlab package to quantify mRNA FISH data. It allows detection and counting **mature mRNA** as well as quantifying the number of **nascent transcripts** at the transcription site. These functions can either be performed cell-by-cell or automatically for a larger population of cells. The main program is FISH_QUANT. This allows accessing and controlling all other functions.

Video tutorials

Several video tutorials illustrating basic workflows in FISH-quant can be downloaded from the website (Installation, how to start, mature mRNA detection, counting nascent transcripts, …).

Example data-sets

Test-data can be downloaded the FISH-quant site. The archive Test_data.zip contains example images together with all quantification results. We will refer to this example with the following box

Example: here references to the different data will be provided.

Typical workflow

Below the typical steps to analyze a FISH experiments are summarized.

Quantification of mature mRNA

1. Define ***outline*** of cells and transcription sites for each image.
2. Finding ***optimized conditions for mature mRNA detection*** by analyzing a small number of cells. Save detection settings. This involves finding good settings for filtering, pre-detection, and Gaussian fitting.
3. Run the automatic ***mature mRNA detection in batch mode***. The detection results can then be checked with the same tools used for individual cells and thresholding parameters to remove false-positives can be adjusted accordingly.
4. Save the results file for individual cells and also the summary file for the mature mRNA counts. The latter contains the summary of mRNA counts for each of the cells in the images.

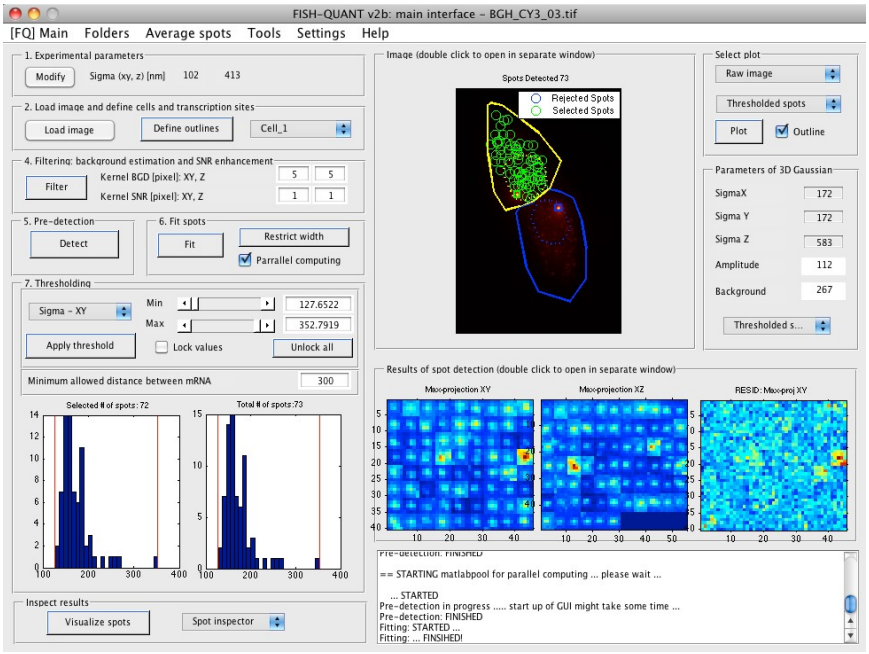
Quantification of the transcription sites

5. The transcription site quantification requires a high quality image of the individual mRNA as well as the distribution of the estimated amplitudes. Both can be obtained after the batch processing for the mature mRNA detection is finished. First, save the summary file of ALL detected spots AFTER thresholding. This file can then be used to define the amplitudes. Then average the images of all mRNAs to obtain a high quality image of the individual mRNA molecules.
6. Set-up the transcription site quantification for a few cells. Save settings for batch processing.
7. Perform transcription site quantification for all cells and save the summary file for the nascent mRNA quantification.

3. Mature mRNA quantification in individual images

This section describes the basic workflow to detect individual mature mRNA molecules in FISH-quant. The screen-shot below shows the main interface of FISH-quant with the results of the analysis of a typical experiment. The following steps are performed for individual images. The detection and quantification settings can then be optimized and used in the batch mode to analyze a large number of images. These steps are also explained in a video tutorial.

The following processing steps have already been performed for the test data. To inspect the results define the relevant folders as described below folder and load the analysis results (see also video tutorial). In this document we will refer to this example data-set at the end of each section.

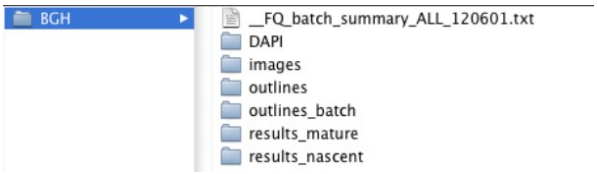


1. Define folders to save data from different processing steps

FISH-quant allows saving and retrieving data from different folders. This is especially useful if a large number of images are processed and different analysis settings are used. If no folder is specified, FISH-quant will assumes that all files are stored together with the image data in ONE folder. If you want to the save the images, outlines, and analysis results in separate folders you can do this with the functions provided in the folders menu. Here four different directories can be specified

- **Root folder**. This directory contains the other directories and will be used as a default if any of the three following directories is not specified.
- **Folder for images**. Contains all the images (raw and filtered).
- **Folder for outlines**. Contains all outline files. Newly generated outline files will be saved in this folder
- **Folder for results**. Contains all results files including the settings files.

Typical folder structure in FISH-quant. Separate folders for images, outlines, outlines generated during the batch processing, and results for mature and nascent mRNA quantification.



Usually the folder for images stays unchanged during the processing. The folders for outlines and results can be changed to save results obtained with different processing settings to different folders. Settings files are always stored together with results files. FISH-quant will always use the folders as defined above to save the different file-types. So even if a different folder is selected in the 'Save as' dialog box FISH-quant will use the defined folder.

You can save the selected folders to a text file with the menu item 'Save folders'. This text file can then be loaded later with the menu item 'Load folders' to get initiate the same folders. The menu item '**Reset folder'** deletes the current folder settings. This can be useful if the folder structure is changed frequently.

2. Define experimental parameters

In this panel the parameters of the experiment can be specified. The user can choose from a number of presets that can be modified. The following values can be defined

- Pixel-size in XY in nanometer.
- Pixel-size in Z in nanometer.
- Refractive index of medium.
- Excitation wavelength of fluorophore.
- Emission wavelength of fluorophore.
- Type of microscope. Currently 'confocal', 'nipkow' and widefield are supported.

These values are used to calculate the theoretical PSF in XY and Z[6]. These values are shown in the panel 'Experimental parameters'. They used for the filtering step and also as initial starting points to fit with the 3D Gaussian.

The default parameters are stored in a text file called *FISH-quant_default_par.txt* in the installation directory of FISH-quant. These values can be adjusted if needed by simply changing the numeric values after the equal sign. Please add no spaces otherwise the numbers cannot be read in correctly.

3. Load image

Two different options are available to load an image (a) load an image by pressing the button 'Load' or the corresponding Menu entry (b) load an outline or results file. Here the associated image will also be loaded. FISH-quant uses the Matlab function `bfopen` to utilize the BioFormats 5.0 Java-library to read a wide variety of image file formats[7].

Example: three raw images can be found in the subfolder *image*.

4. Outline of cells and transcription sites

This opens a separate user interface that allows defining the outline of the cell(s) and their associated transcription-site(s). The detection and quantification of mature mRNA will be only performed for the areas corresponding to the defined cells. The area occupied by transcription sites will be excluded.

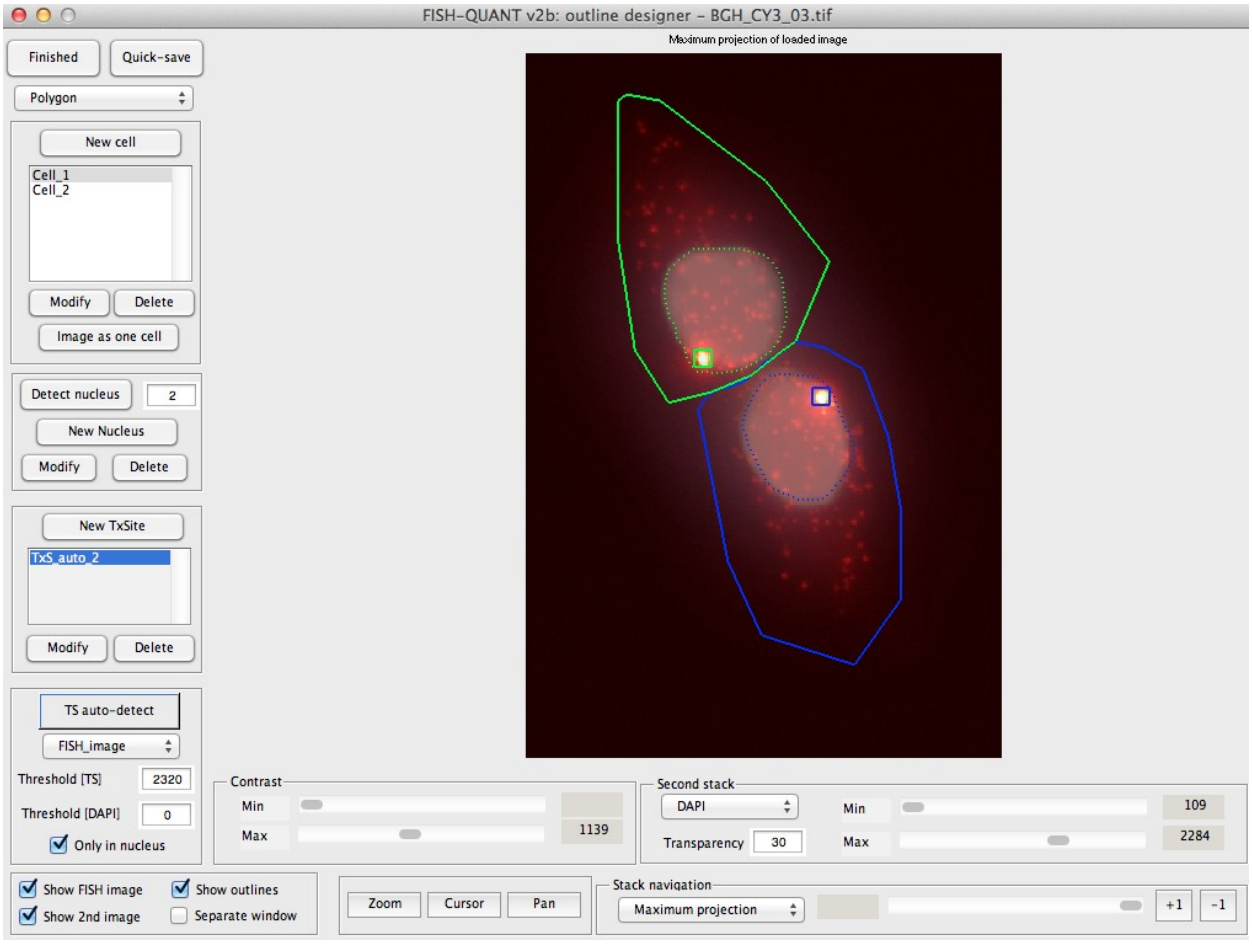
Additionally, a nucleus can be defined for each cell. This can be done either manually or by an automated detection based on the DAPI signal (see below).

Manual definition of cells and transcription sites

Outlines can be defined as freehand region, polygons, rectangles, or ellipses. These geometries can be mixed in the same image. Selected the desired geometry from the drop-down menu. To **define a new cell** press the button 'New cell' or press the keyboard button **C** and outline the cell. For a freehand region, simply stop following the outline when done. For polygon, rectangle and ellipses d**ouble-click to finish drawing**. The outline of the cell will then be displayed with a solid, green line.

To *modify the outline a cell*, select it from list and press 'Modify'; to delete press 'Delete cell'. **Transcription sites** can be defined similarly. Please note that transcription site can only be defined within an already defined cell. Draw the transcription site and FISH-quant will automatically assign the site to the corresponding cell. When finished, press button 'Finished' to return to main interface. The GUI can be opened again to change outlines if necessary.

Important: when defining a new cell or a new transcription site the cursor will change from an arrow to a cross. Depending on the used computer this might take a few seconds; please wait until the cursor changes appearance before continuing.



The FISH image can be either displayed as a XY-maximum intensity projection or as a z-stack where a certain plane can be selected. Please note that the defined regions extend through the entire z-stack. It is

possible to **zoom and pan** make the selection of individual cells easier. The **contrast** of the displayed image can be changed with the sliders at the bottom interface. Contrasts can **NOT** be changed while drawing a region!

Saving and loading of outlines

The outline of the cell, the transcription site(s), and if specified the nuclei can be **saved** in a text file (see Appendix 1). This file can also be loaded again and is used for batch-processing (see below). This files contains not only the outlines, but also the name of the images and other experimental parameters. FISH-quant allows saving also a reference to files with DAPI signal or an independent label of the transcription site (see below). If such files are present, FISH-quant will ask if they should be loaded.

Quick-save: when pressing this button FISH-quant proposes a default file-name to save the outline. This avoid going through the menu to save the outline.

Other features

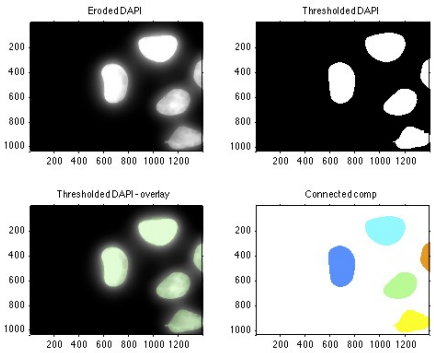
- If no outline is defined the subsequent detection and quantification steps will be performed on the entire image. Alternatively the entire image can be selected as one cell with the option 'Whole image as cell'. Then transcription sites can be defined. This is a useful feature to process for instance images of fluorescent beads. The 'transcription sites' could be region excluded from the analysis.
- By selecting the 'Image in separate window' option the image will be shown in a separate figure. It is possible to **zoom into specific regions** for a more accurate selection. After selecting 'New cell' or 'New TxSite' press on the zoom icon and the cursor will turn into a magnification glass. Then enlarge the region of interest. Revert the cursor to the cross for outline selection by pressing on the zoom icon again. When the selection is finished the main interface for the outline selection will be shown. If a new cell or transcription site is defined, the same image as before (with the same zoom settings will be shown).

[Optional] Definition and detection of nuclei

FISH-quant also allows defining a nucleus for each cell. The nuclei can be used to limit the detection of the transcription sites (below) and can be used to distinguish the localization of mRNA in the nucleus and the cytoplasm. The definition of the nucleus is in XY and extends throughout the entire z-stack; additionally each nucleus has to be inside of a cell. The outlines of the nuclei can be best defined with an additional 3D stack containing a nuclear stain, e.g. DAPI. The DAPI image can be loaded from the menu *Load Images* > *Load DAPI* or with the short-cut **CTRL-D (WIN) or CMD-D (MAC)**. By default the DAPI image will be shown as an overlay with the FISH image. It's transparency and contrast can be changed with the corresponding controls.

The outline of a nucleus can be defined either manually (as for the cells) or automatically. For the automated detection, an intensity threshold has to be defined. FISH-quant provides an automatically calculated default value (value between 0 and 100, numbers with commas ARE allowed). Press '*Detect nucleus*' to apply this threshold (Figure below shows a plot that will be displayed during this detection). Nuclei will be automatically detected and assigned to the corresponding cells. When selecting the option 'Only in current cell', the detected nuclei will only be assigned to the currently selected cell. However, this methods works only for nuclei that are well separated and for nuclei that are too close it will fail. Here the user can manually outline the nuclei. This option is also available via the keyboard button **N**.

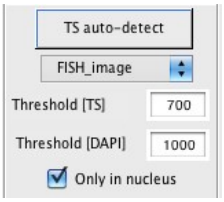
FISH-quant will show a plot illustrating the different steps in this process (show on the right). If the defined threshold doesn't provide good results change it until DAPI signal is satisfyingly thresholded.



[Optional] Automated detection of transcription sites based on their intensity

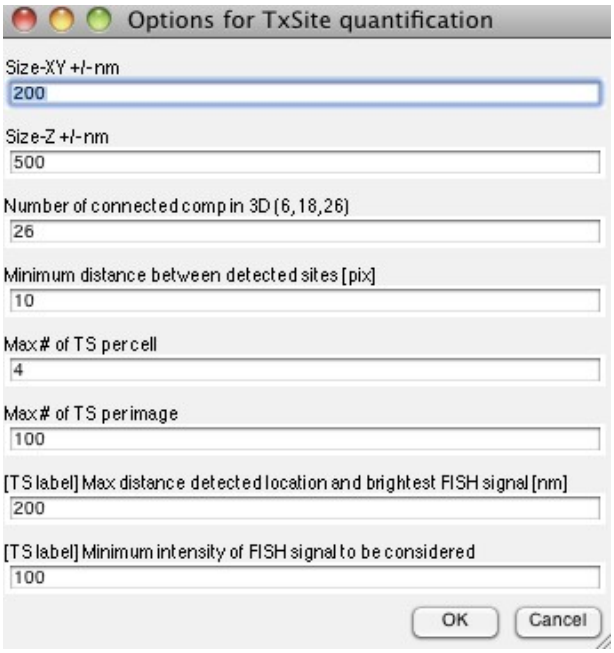
Transcription sites can also be automatically detected. FISH-quant identifies regions that are above a specified intensity threshold and assigns them to the corresponding cells. This threshold is chosen such that it separates mature mRNA from transcription sites. To get a feeling for the intensity values you can use the Cursor tool to inspect the intensity values.

Then use this intensity value in the text box for the minimum intensity (700 in the example on the right), choose 'FISH-image' and press 'TS auto-detect' to perform the detection. The plot will then be updated and the detected sites shown.



For lower threshold values this detection can lead to false-positives. To **reduce the number of these false detections the DAPI signal, or any other nuclear stain, can be used**. The user can either restrict the detection to the nucleus as determined by the DAPI (enabled in the example above). Alternatively, the user can specify a minimum intensity value a detected transcription site has to have in the DAPI channel. The DAPI signal inside a nucleus is usually much higher than in the cytoplasm so normally even relatively conservative thresholds provide satisfying results. This option is especially practical if the outlines of nuclei are not defined.

Several options for the detection can be changed from the menu *Options> TxSite detection*.



- The cropping ***size*** around each detected transcription site. This is expressed relative to the brightest voxel of the detected site. These values should be chosen such that the transcription site fit into the cropping range (including some background).
- Sites are detected as **connected components** with the Matlab function `bwlabeln` (See also Matlab help file for more information about connected components). The three-dimensional connectivity can be changed. Howver, the default value should work in most cases.
- The allowed **minimum distance** between the centers of two detected transcription sites (in pixel) can be defined. This is important if relatively large transcription sites are present and avoids that they are detected as two separate sites
- The program also minimizes the **over-detection of sites**. First, the maximum total number of sites per image can be restricted. If more sites are detected the detection will be stopped. This usually means that the used intensity threshold was too low. Second, the maximum number of detected sites per cell can be restricted. If the detected number of sites exceeds this limit, FISH-quant will only use the brightest ones.
- The last parameters are only relevant if the detection of the transcription site is performed based on an image with an **independent label**, especially LacI (see below). The LacI repeats are not always directly next to the gene. Therefore the LacI signal and transcription site as detected by FISH might not completely overlap. FISH-quant therefore allows a maximum distance offset between the detected LacI location and the FISH signal. The location of the transcription site is first detected with the LacI signal. Then FISH-quant considers an area defined by this offset surrounding the detected location in the (raw) FISH image. The program looks for the brightest voxel in this area. The detected location will then be corrected if this voxel is brighter as a defined intensity threshold. This last threshold is necessary to avoid that the detected location is corrected if not transcription site is present. Therefore the intensity value should be set such that it reflects at least the brightness of an individual transcript.

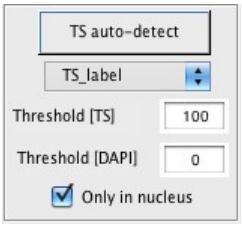
[Optional] Automated detection of transcription sites with independent label

The detection method based on the intensity of the transcription sites works best for bright and strongly transcribing sites. For weakly transcribing sites this methods fails, especially if only one transcript is attached to the site. Transcription sites can then only be reliably detected by independently tagging the transcription site with a second marker. Here, several experimental approaches have been developed. First, in the LacI tagging approach[8] the lac repressor (LacI) is tagged with GFP and binds to arrays of lac operator sequences inserted at close to the transcription sites on the chromosome. Second, mRNA FISH can be performed with probes designed against the intron of the studied genes. Most transcripts are spiced co-transcriptionally so a FISH signal is only obtained for transcription sites. Third, DNA FISH can be performed against the target genes.

Ultimately, each of these methods produces a second 3D stack where the transcription sites are marked independently. In FISH-quant, these additional images can be used automatically detect transcription sites also in the absence of FISH signal.

We recommend **filtering** these images before using them for automated detection. We provide a small GUI to perform this task for a large number of images (See section BATCH FILTERING for more details). The filtered image can then be loaded from the menu *Load Images* > *Load TS-label* or with the short-cut **CTRL-T (WIN) or CMD-T (MAC)**. The image can then be displayed by selecting it from the pull-down menu in the panel 'Second stack'. It's transparency and contrast can be changed with the corresponding controls. It can also be displayed alone by enabling only 'Show 2nd image'. Then transcription sites can again be detected on an intensity threshold.

The Cursor tool can be used to find a good threshold value. Then 'TS_label' has to be selected (see right). The detection can again be restricted to the nucleus or a second intensity threshold for the DAPI signal can be defined.



Additional settings can be defined in the general settings for TxSite detection (see preceding section). They allow correcting for a possible offset between the TxSite label and the actual position of the site.

For a **more detailed step-by-step description** see also the separate help file FQ_TS_detection_LacI.pdf that can be found in the documentation subdirectory of FISH-quant.

[Optional] Save settings for automated TxSite detection

The settings used for the automated detection of transcription sites can then be save in a simple file from the menu [*FQ] outline* > *Save settings for TS detection*. These settings can then be used to detect transcription sites in batch processing (see section *Detection of transcription sites in batch processing*).

5. Defining outlines for batch processing

In FISH-quant a large number of images can be processed automatically with the batch processing tool. The outlines for each image have to be drawn manually before. This can be done with the same interface described above but loaded from Menu *Tools* > *Outline designer*. Now the button at the top is called 'Open Image' and can be used to load the different images (or alternatively the shortcut **CRTL-F (WIN) or CMD-F (MAC)**). Then their outlines can be defined and saved.

A useful option when working with DAPI or TS-label image is the possibility to **define default names**. When the Outline-designer is called separately, the user will be asked if he wants to use default names the first time a DAPI or TS-label image is opened. Press YES to define some parameters for this option, NO for not using this option. After pressing YES, **unique** identifiers for the FISH, DAPI, (and if used) TS-label have to be specified. For instance, this could be *_CY3_* for FISH and *_DAPI_* for DAPI. FISH-quant will then look in the file-name of the FISH images for the identifier for FISH and replace it with the identifier for DAPI. This new name will then be used as a default name to open the DAPI image. These identifiers have to be defined only once (but they will be lost as soon as the Outline designer will be closed). Then the DAPI images can be automatically selected by pressing CTRL-D and loaded by confirming by pressing *enter*. The same option applies to TS-label images.

Another interesting way to speed-up this step is 'List directory content' tool which is accessible from the main FISH-quant menu (Tools). Here you can navigate to the folder containing the images. When clicking on an (FISH) image, it will be automatically opened in the outline

designer. For the first image, the default names have to be specified as above. For the subsequent, they will be remembered. Like this FISH images can be easily opened in the Outline designer, then by pressing CTRL-D + enter the corresponding DAPI image is loaded.

Example: the outline files for the three raw images can be found in the subfolder *outlines*. They contain no reference to the filtered image (see below). The folder *outlines_batch* contains outline files that were generated automatically with the batch-processing tool and contain a reference to filtered image (see section on batch processing for more detail).

6. Creating outlines for a second channel

For certain biological question, smFISH is performed in different colors for instance to target different mRNA species in the same cell. To facilitate the analysis of these data, FISH-quant provides several options to generate outline files for a second channel based on existing outline files. These new outline files are in general generating by (i) renaming the actual outline file to reflect the new channel (ii) replacing the referenced image within the outline file. Please note, that we do not change the emission and excitation wavelength within the outline file for the first two options since this has negligible impact on the analysis.

Image-by-Image

You can define in FISH-quant the outlines for one given color for one or multiple images and save each of them. Then you can open images in the other color one-by-one and select from the menu *[FQ] Main* > *Load* > *Outline from another color*. This will overlay the selected outline on the currently opened image. Then save the outline which will now generate an outline file with the currently opened image and the loaded cellular outlines.

Batch-mode

From the tools menu, you can open the module 'Create outlines 2nd channel'. You first have to specify unique strings for each color, e.g. *CY3* and *CY5*. By replacing the string of the first color in the file-name of the outlines/images by the string of the second color, you have to obtain the file-name of the second color. For instance, let's assume the outlines in the first color are called *Test_pXY_CY3_outline.txt*, where *pXY* are the different images (*p01*, *p02*, ..); and the outlines of the second color are called *Test_pXY_CY5_outline.txt*. Then replacing CY3 by CY5 will yield - as required -for each outline of the first color the outline of the second color. The same has to hold true for the image names that are referenced within these outlines files. Then press *Create*. Here you have to select all outlines in the first color that you want to convert. FQ will then open each outline and will (a) change the name of the referenced image in the outline (b) save this outline file under the new file-name. These files will be stored in a subfolder with the name of the unique string of the second color, e.g. CY5 in our example.

Script for more than one additional channel

If you want to generate outlines for even more colors, then you can use the script **SCR_outlines_create_v1**. You have to first specify in a **tab**-delimited text file all the colors that you would like to generate. The first row contains the unique string of the first color, e.g. CY3. The following rows define one color each. First element is the unique string, e.g. CY5; 2nd element is the excitation wavelength in nm, 3rd element the emission wavelength. In the script you first define all the outline files that you would like to process, then you can load the text file defining the different colors. The script will then generate the outlines for the different colors, and save them in a separate folder indicating the unique string of this color.

7. Filtering

The pre-detection is performed on a filtered image. FISH-quant provides two different filters

(1) **Dual Gaussian**[9],

(2) **Laplacian of Gaussian** (LoG)[10].

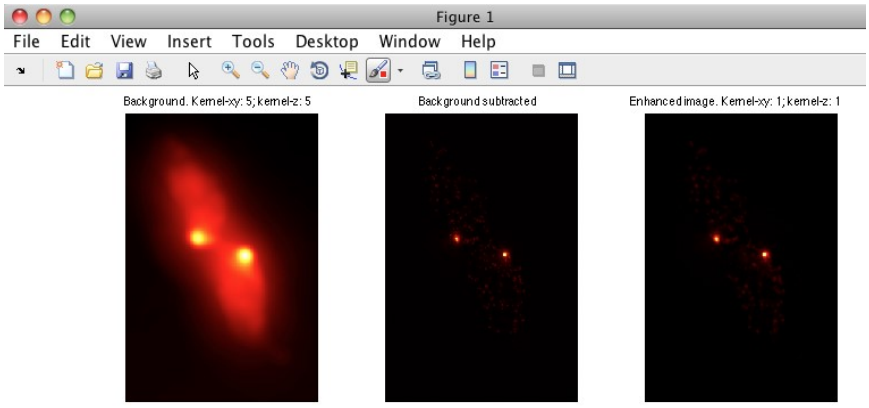
Both filter will reduce the background contribution and increase the signal of the individual spots. We found that their performance is quite comparable, with sometimes a slight better performance of the LoG filter. Below we only detail the Dual Gaussian filter; similar considerations apply also for the LoG filter.

In some cases, **you might not want to filter at all**, in this case set all filter parameters to 0.

In the **Dual-Gaussian filter**, a two-step filtering process reduces noise and enhances the signal-to-noise ratio (SNR). First, the image is filtered with a **large Gaussian Kernel** to blur the image. This yields a good approximation of the background and will be subtracted from the raw image. The resulting image is then filtered with a **small Gaussian Kernel** to enhance small features. The **size of these two kernels** is expressed in pixels (decimal values are also possible) and can be changed in the filtering panel.

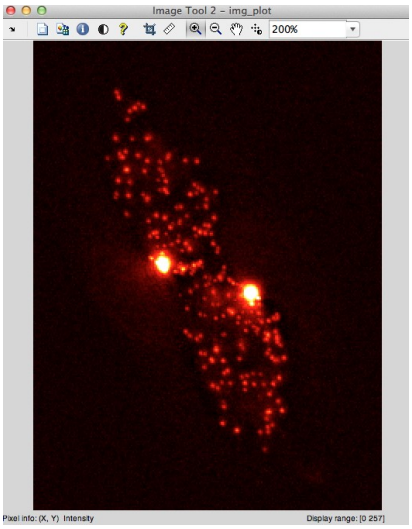
For images with spots that are close together **a relatively small Kernel for XY (<1)** can help to detect spots that are close. Here you have to try different values and inspect the results of the pre-detection for the different settings (by looking how the detection in a relatively dense region worked).

It is possible to define **different Kernel sizes for 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. When the filtering is done, FISH-quant displays a plot (below). This image is especially useful to judge the background image.



It might, however, be hard to see the individual spots if bright features such as transcription sites are present. In this case the filtered image can be displayed in a special window. In the ‘Select plot’ panel in the upper right corner of the main interface select ‘Filtered image’ and disable ‘Outline’. Then simply double click on the image. This will open a separate figure displaying the filtered image.

In the menu bar are controls to zoom, pan, and change the contrast (More information can be found in the MATLAB help when searching 'Image Tool Overview'). When moving the cursor over the image MATLAB will display in the lower left corner the position and intensity of the pixel that is under the cursor. This allows inspecting the intensity range of the individual spots in the filtered image and judge changes of different values of the filtering Kernels.



The default settings for the filtering usually provide good results. We still advice to test other settings, especially if new imaging conditions have been used. This can be done by simply filtering the images with different combinations of the Kernel sizes and saving the resulting images (see next paragraph). These images can then be compared in ImageJ to select the optimal size of the either Kernel.

The filtered image can be saved as a stacked TIF file (Menu: Load/Save > Save > Filtered Image). The file-name of the filtered image will then also be stored in the outline file and in the results file (Appendix 1 and 3). This is saves time if the same data sets will be re-analyzed since the already filtered image can be used. It is, however, necessary to FIRST save the filtered image and then save the outline definition file. Alternatively the file-name of the filtered image can manually be added (see Appendix 1).

IMPORTANT: it is NOT necessary to filter each image manually. Once good settings have been determined, the filtering can be done automatically with the batch-processing module (see corresponding section for details). FISH-quant will filter the images and generates new outline files with a reference to these image.

Example: filtered images of the example data are in the subfolder *image*. They were generated automatically with the batch-processing tool (see below).

OPTIONAL: it is also possible to load already existing filtered images from the menu. These images can for instance be obtained from a deconvolution software. Then the described filtering step has to be performed. See also section on batch processing how these images can be loaded automatically when processing many outline files.

8. Pre-detection

Pre-detection **is performed on the filtered image** and settings are defined in a separate graphical interface (see screen-shot below). The definition of these parameters is based on the currently selected cell from the main interface. After the pre-detection parameters are determined, the pre-detection will be applied to all other cells in the image. We therefore recommend selecting a cell with a large number of spots in the main interface before starting the pre-detection.

The pre-detection involves two main steps.

1. **Possible spot candidates** are identified based on an intensity threshold.

2. A **quality score** is calculated for each candidate and is used as a thresholding parameter.

The user has to define the threshold for both steps. **In general** it is advisable to allow some false positive detection. Bad spot can then later be excluded by the various estimated fitting parameters. So in general a relatively modest thresholding strategy should be applied.

The user can further choose in **which part of the cell the detection will be performed**. By default the entire cell is considered. If the outline of a nucleus is defined, the user can choose to perform the detection only in the cytoplasm or only in the nucleus. This can be useful if cells with a strong nuclear background are analyzed. Here two different settings can be defined to analyze nucleus and cytoplasm separately. For cells where no nucleus is defined, the detection will be performed in the entire cell, independently of the selection.

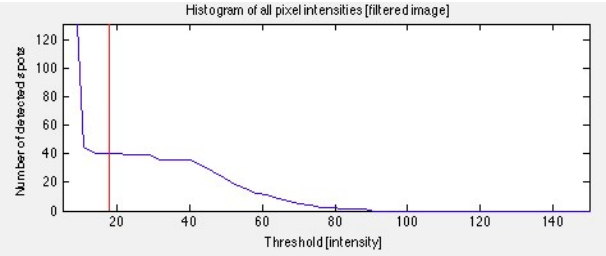
Inspection of filtered image

The basic concept of the pre-detection is to define an intensity threshold that distinguishes actual spots from background. We therefore recommend beginning with a visual inspection of the filtered image as described above to get a feeling for intensity range of individual spots in the filtered image.

Setting the range of intensity thresholds that will be tested

The pre-detection GUI then allows specifying the intensity detection thresholds.

To facilitate the definition of this threshold FISH-quant will display a plot of the number of detected spots as function of different intensity thresholds as shown on the right. This plot will be explained in more details next.

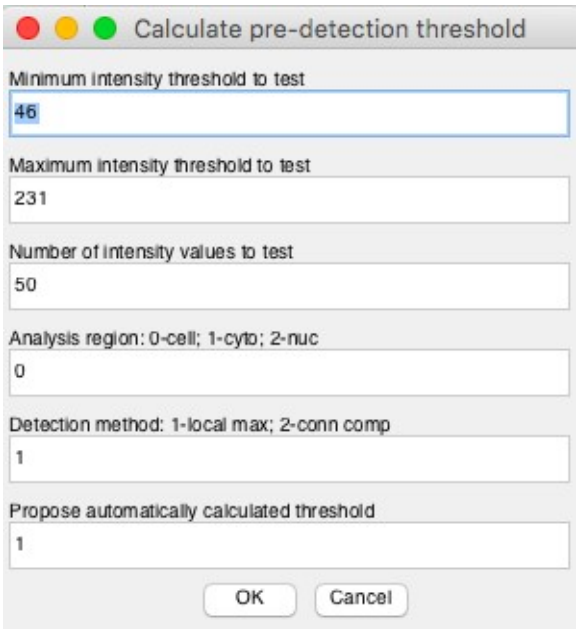


Defining initial parameters for detection

Before the GUI is displayed a dialog is shown that allows specifying the range of tested intensities and where the detection will be performed.

The first time a prediction is performed on an image, this range is calculated automatically with the maximum value being 3 times the 99% quantile, and minimum value being 5% of the maximum value. These values can be further adjusted if needed. Usually only the maximum intensity value has to be adjusted.

- **Minimum** intensity corresponding to background; should be smaller than the dimmest spots.
- **Maximum** intensity corresponding to intensities brighter than the actual individual spots, as a rule of thumb 1.5 brighter than the brightest spot.
- **Number of tested intensity** values. The minimum difference between tested intensity values is 1. The number of tested values will be adapted if the specified value would result in a finer spacing.



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 in the GUI afterwards.

Then you can specify the **pre-detection method**: **1** for local maximum [default] or **2** for connected components.

With the last option you can choose if you want FISH-quant to propose an automatically calculated threshold (see below for more details).

These parameters will then be used to perform a preliminary pre-detection for the specified range of intensity value. This usually ***causes a delay*** of several seconds before the GUI is displayed. Parameters can also be changed in the GUI if the specified range was not good (see below).

Sub-region around each spot for detection and fitting

For each pre-detected spot FISH-quant extracts a **sub-region** around the detected center. This smaller image will then be used for all subsequent processing steps (calculation of the quality scores and fitting of the 3D Gaussian). The size of this region is specified in the upper part of the GUI and expressed +/- number of pixel with respect to the identified spot location. By default this value is set to twice the size of the theoretical PSF. Larger or smaller values might be desirable depending on the experimental setup. Please also check the results of the Gaussian fitting if the defined sub-region captures the relevant part of the signal. By default only spots for which the sub-region can be extracted will be considered in the analysis. This will exclude spots that are close to the first or last focal plane. Selecting the '*Allow smaller region for detection in Z*' option disables this setting. Now all spots independently of their z-position are considered.

Pre-detection of spots

The **identification of spot candidates** can then be performed with two different functions. The first method uses **non-maximal suppression** (`nonMaxSupr` function of Piotr's Matlab toolbox[1]). This method identifies pixel location and values of local maximums - that is a location is returned only if it has a value greater or equal to all pixels in the surrounding window. The second method is based on **connected components** (Matlab function `bwconncomp`[10]). Here candidates are identified based binary image generated with the specified threshold. Connected regions in this binary image are considered to be spots and used for further analysis. In either method the minimum intensity a spot must have has to be specified. When changing the pre-detection method some calculations have to be done. This might take a

few seconds. This calculation is, however, done only once for a give set of thresholds and afterwards you can simply switch between the two pre-detection methods.

The **plot** in this panel shows the number of detected spots as a function of this intensity threshold. By moving the slider (or editing the text box) the threshold intensity can be set. The value should be set such that all spots are considered without considering too much background. Usually a plateau region can be seen for an optimal threshold. We usually place the intensity threshold towards the left part of the plateau, which will lead to a slight over-detection, but the subsequent steps will remove false-positive detections.

Please note that the tested range of thresholds might not be optimal for your data. The plot is generated with a **pre-defined set of intensity values as described above**. You have to manual change this range by editing the values in the corresponding text boxes. To apply the changes press 'Apply'. This will again take a few seconds.

When using the method based on the local maximum the size of the sub-region will also be used as the area for where the local maximum is detected (see function nonMaxSupr for details, the subregion is used as the suppression window dimensions). For certain image conditions this might, however, not yield good results, e.g. when images are obtained with very fine sampling in z and small pixel size in xy. Here the user can define the suppression window dimension separately by enabling 'Different region to detect position' and setting the size accordingly. Usually this size is then set to values smaller than the sub-region used for the sub-sequent analysis.

Automatic threshold determination

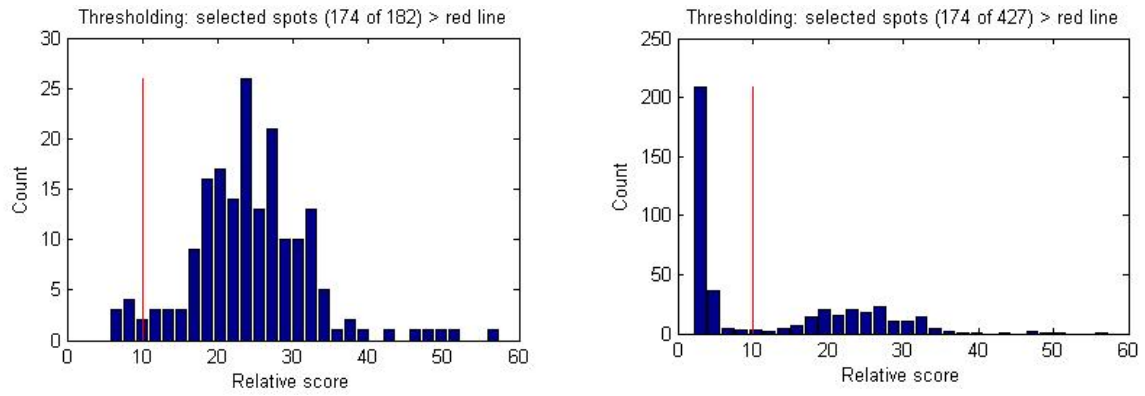
You can enable this option when starting the pre-detection. FISH-quant attempts then to automatically calculate the detection intensity threshold. To this end, FISH-quant uses the number of estimated mRNAs as a function of different tested intensity thresholds. First, we calculate the gradient of this curve, i.e. value that will be small for the flat parts. We then calculated the inverse absolute value of the gradient, i.e. values will now be large for flat areas. Ideally, this happens at the plateau corresponding to the best detection threshold, but it will also occur towards the end where the curve goes to 0. Therefore we multiply this value with the corresponding number of detected spots at the underlying intensity threshold. In an ideal case, this will be a maximum value around the best detection threshold. However, if the plateau is not perfectly flat, this methods starts to fail. We found that in same cases, down-sampling the data helps to get better estimates. FISH-quant therefore calculates the intensity threshold on all data, and on data downsampled by a factor of 3, and suggests their mean value as a detection threshold.

Quality score of spots [optional]

When the pre-detection is finished a **histogram of the quality score** is shown and the user can specify a cut-off. These quality score are an additional check that allows distinguishing noise from real spots. Two scores are implemented

- ***Standard deviation*** of all pixels in the sub-region. Works robust also for dimmer spots
- ***Curvature***. 3D curvature [11]. Works best for bright spots.

In the title of the histogram the total number of spots and the currently consider number of spots are shown. Good spots populate the right part of the histogram, whereas bad spots can be found on the left.



(Left) Pre-detection worked well and threshold is chosen such that entire distribution is selected. (Right) Pre-detection yielded also many false-positives. Threshold should be chosen to separate the two distributions.

- The histogram has usually only one peak (plot above, left panel) if the pre-detection worked well and no false-positives were detected. Here the quality-score can be set (close) to zero.

- If the detection also yielded false-positives the histogram (can) show two peaks and the threshold should be chosen such that the two peaks are well separated (plot above, right panel).

- We found that such a **clear separation is often not the case**, in which case it is preferable to simply adjust the pre-detection value as good as possible and omit the quality score.

Press '**Finish'** to finish the pre-detection. The determined pre-detection settings will then be applied to all other cells in the image.

9. Semi-automated calculation of detection threshold

We implemented a simple Matlab user interface (*FQ_detect*) to facilitate the determination of the detection threshold threshold. In this tool, a FQ outline file specifying all cells in a given image can be opened. It then crops a pre-defined number of cells and applies the specified filter. Finally, it proposes an automatically calculated threshold based on Otsu thresholding.

We identified empirically three possible cases that require different computation of the threshold. (1) Two class Otsu works for simple cases with isolated mRNA molecules above uniform background, and the determined threshold be used as detection threshold; (2) the presence of bright structures (such as foci) requires a three classes Otsu, and the threshold separating the middle class from the dimmer class (background) is used as a detection threshold. The presence of bright structures is tested by verifying if the value provided by the two class Otsu is greater than 10 times the median of all pixels. (3) If the background shows dim patches, the results of a three classes Otsu have to considered and the threshold separating the brightest class from the middle class be used as an intensity threshold. The presence of dim patches is tested by verifying is the threshold of a two class Otsu is smaller than 2 times the median of all pixels. We found that this simple approach works well for images with reasonable quality, and detection thresholds can be set fast for a large number of experiments.

1) You can specify the maximum number of cells per image that will be considered in this analysis. By default, the tool uses the number of CPUs available via the parallel computing option. However, this number can also be increased.
2) You have to specify the folder containing the images
3) Specify the filtering settings you want to use.

4) Load the outline file that you want to process. Note that if you change the filtering settings / detection settings, you have to reload the outline file.
5) The program will display the first cell with the automatically determined detection threshold. You can enable / disable the display of the detected spots with the corresponding button.
    - You can select other cells from the listbox on the left side.
    - You can change the display contrast of the cell with the corresponding button. The same contrast settings will be applied for other cells as well.
    - You can manually change the detection threshold and re-perform the detection.
6) Once satisfied with the results, you can save the settings with the corresponding button. They will be saved in the same folder as the opened outline file.

10. Fit spots with 3D Gaussian

The pre-detected spots will be fit with a 3D Gaussian function (Appendix 9). This step uses automatically **parallel computing** (if available)**, which** reduces the processing time on computers with multiple cores[1]. The following parameters are estimated:

- Width of the Gaussian in XY
- Width of the Gaussian in Z
- Position in X, Y, Z
- Amplitude
- Local background

As soon as the fitting process is finished a number plots in the main interface are updated. There are three smaller plots under the plot window. The first two panels show a maximum projection of all detected spots in XY and XZ is shown. The last panel shows the maximum projection of the residuals in XY. The spots can then be thresholded (see next section 'thresholding') and the results visualized (see section on visualization). Fits can also be re-done with a restricted range for the width of the 3D Gaussian (see below).

11. [Optional] Restrict number of spots per cell that will be fit

Fitting spots that have spots in close proximity can cause problems such as erroneous estimation of the position of the spot and its width. This is caused when a neighboring spot contributes signal to the subregion of the currently fit spot. Specifically, the sigma's tend to get larger in this cause and the estimated position moves towards this additional signal. This causes problems when using sigma's as a thresholding parameters (see below) because a good spot will be removed (treated as a false-negative).

One way to minimize this problem is to reduce the size of the region cropped around each spot. However, this might not work for really dense cells. Here FQ provides the possibility to fit spots only in cells with a user defined maximum number of spots. This option is available when pressing the button 'Restrict # of fitted spots'. A value of -1 means that all spots in all cells will be fit, a value of >= 0 means that only in cells with fewer spots than the specified number a fit will be performed. For example, when setting this value to 20, FQ will fit spots only in cells having fewer than 20 spots. It will then assign 0 as a value for the different fitting parameters; expect the positions where it will use the pre-detected position. In the results file, the missing values will be indicated as 'NaN' (short for not a number).

[1] If you don't want to automatically use parallel computing, you can deactivate this behavior in the preferences for parallel computing by disabling the automatic start of parallel pools.

Application example: a gene is very heterogeneously expressed. The pre-detection works relatively well for all expression levels, however, the fit is causing problems. Additionally, the TS should be quantified. Here, the results of the pre-detection could be used as an estimate of the number of mRNA molecules per cell. The TS quantification could be performed based on the mRNA signal extracted from cells with a low mRNA copy number. To achieve this the fitting can be restricted to cells with fewer than a certain number of mRNA molecules. By not setting any threshold after the fit was performed, all mRNA can be counted. Then a threshold for the sigmas can be applied (slightly larger than 0) to remove all the spots that were not fitted. With this selection the averaged mRNA image and the spot summary can be generated and used for the TS quantification (see section on TS quantification for more details).

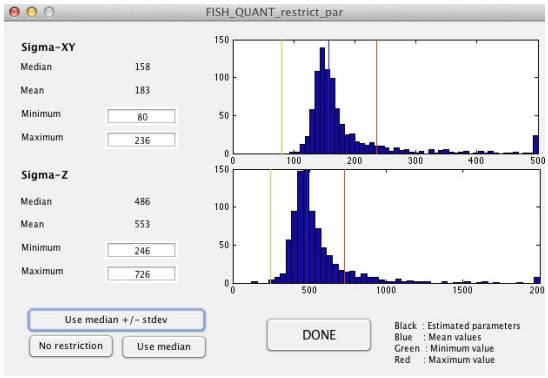
12. [Optional] Restrict allowed range for width of 3D PSF

The range of the allowed sigma's for the 3D Gaussian can be restricted. Press 'Restrict width' to open the corresponding user interface shown below. In this GUI (see plot on the right) the allowed range for the width of the Gaussian in XY, and Z can be restricted. Restricting the parameters to a reasonable range can help to reduce the noise fluctuations in the quantification. Please note that the minimum and maximum values have to be different. To fix the fitting parameters to a certain value assign this value to the minimum and set the maximum to this value +1. A number of pre-defined ranges can be set by pressing the corresponding button.

- ***'No restriction'*** gives a very wide range to allow a free parameter estimation.
- ***'Use default'*** give the default setting which are median value +/- standard.
- ***'Use median +/- stdev'*** gives the median +/- standard deviation for all parameters
- ***'Use median'*** fixes the range to the respective mean value.

Press DONE to set the define range and return to the main interface.

The histograms show the results of the last round of quantification. The blue line indicates the media value; the green and red line the minimum and maximum allowed value, respectively.



13. Thresholding

In this panel, the fitted spots can be thresholded based on the estimated parameters. The parameters can be selected from the pop-up menu and two histograms will be displayed. The histogram on the right shows the value for all spots, the histogram on the left only the values after thresholding. The values are for the currently selected cell. Minimum and maximum values for this parameter can then be specified and their values are indicated in the histograms by vertical lines. The thresholds are applied by pressing the button 'Apply threshold'. To save these thresholds enable '**Lock values'**. Then the next parameter can be inspected. The option 'Unlock all' removes all applied thresholds. The same thresholding parameters will be applied for all analyzed cells.

Detected spots can be very close to each other and even overlap. This happens usually when the pre-detection yielded a large number of false positives. To avoid multiple detections of spots like these the user can define a minimum distance between spots. This value is specified in the field 'Minimum allowed distance between mRNA' (the default values is the pixel-size in z). FISH-quant then calculates the pairwise distance between all spots. Then spot-pairs (or triplicates, …) that have less than this value are identified. FISH-quant then removes all expect one spot with the highest intensity in the pre-detection.

In the panel 'Parameters of the 3D Gaussian' the averaged values for each of the estimated parameters is shown. The pull-down menu allows selecting either the average of all spots, all thresholded spots, or (if calculated) the results of fitting the averaged spot. For each new selection in the pull-down menu the results will also be shown in the workspace together with their standard deviation.

Note: a matrix with the parameters of the thresholded spots is available as the global variable `spots_fit_th` and can be used in other Matlab scripts. This matrix has the same columns as the part of the results file describing the detected spots (see Appendix 3). The matrix will be updated after each thresholding step.

14. Inspection of detected spots

FISH_QUANT allows a detailed inspection of the detected spots. Three different visualization options are available

- **Spot-inspector.** Matlab GUI that allows inspecting individual spots and their estimated parameters.
- **3D-rendering**. Uses the matVTK package to show the detected spots in rendered 3D plot.

Spot-inspector

When selecting the Spot-inspector a new GUI opens (see screenshot below). The image can be shown either as a maximum intensity projection or as a z-stack where you can scroll through the slides. Furthermore, either the raw image or the filtered image can be shown. Lastly, you can select if all detected spots should be shown or only the spots that past the thresholding. The GUI can be closed by the 'Close Window' button in the upper left corner.

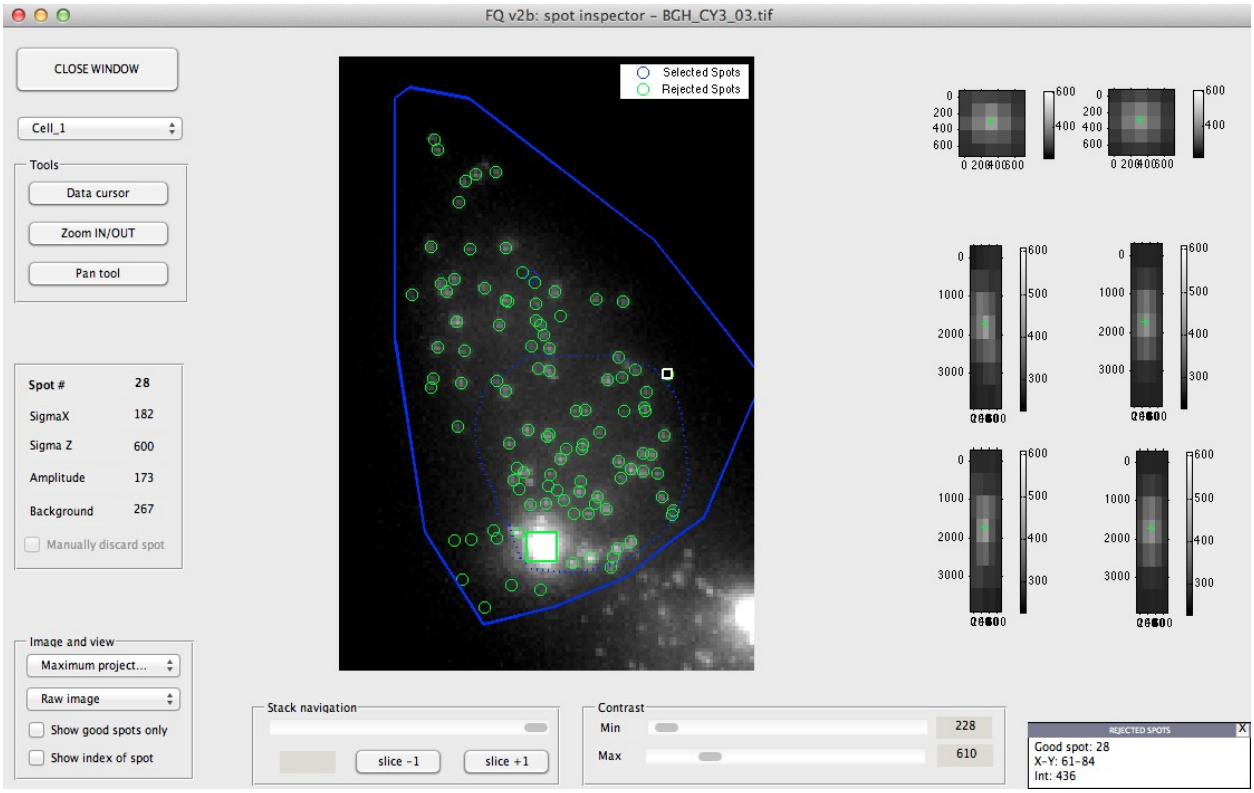
- The **detected spots** can be inspected individually with the **Data cursor**. When selecting one spots the sub-region of this spots as used in FISH-quant will be shown in the panel windows on the right side of the interface together with the best fit. Additionally, the corresponding fitting parameters will be shown on the left side of the interface.

- The **Region inspector** opens the Matlab region inspector which allows a more detailed inspection of the image data.

- **Zoom In/Out** allows zooming the image. This can either be done with the mouse wheel or by pressing the left mouse button to zoom-in and shift+mouse to zoom out.

- **Pan** activates the pan mode that allows moving view of a (zoomed) graph up and down as well as left and right.

- The **index of each spot** can be displayed next to the spot in the image by enabling the option 'Show index of spots'.



This tool can be **opened in different ways**.

- When opening it from the main FISH-quant the currently processed data-set will be shown.
- The 'spot inspector' can be opened from the Tools menu in the main interface to inspect already saved results file. Use the folder symbol in the ribbon to load a file. The fitted images of the individual spots are not shown. Here, it is possible **to manually discard spots**. For this, you first have to select a spot with the 'Data cursor' and then enabling 'Manually discard spot'. If you discarded spots and you want to keep this selection, you have to save the results file (by pressing on the disc item in the ribbon)!

  Note 1: it is important to define the folder structure – otherwise this option will not work.

  Note 2: the identical interface will be opened when opening files with the tool 'List directory content'
- **The tool 'List directory content'** can be used to navigate to the save results of the spot detection. When clicking on the results files, they will be opened in the Spot inspector.

15. Saving and loading the quantification results

Two important files summarizing the mature mRNA quantification can be saved from the menu *[FQ MAIN] > Save* (see Appendix 3 & 4) or with the short-cut **CTRL-S**.

1. **Results of the actual spots detection** in a simple text file. This file contains all the estimated parameters for the detected spots of one cell. In addition it repeats the information contained in the outline-definition file (see Appendix 3). Either all spots can be save ('Detected spots [All]') or only the spots that passed the thresholding ('Detected spots [Thresholded]').

2. The **detection settings** are saved as a text file and can then be loaded again.

These files can also be loaded again from *[FQ MAIN]* > *Load.* Before doing this the folder structure has to be accordingly updated (see above).

Example: the results file for the spot detection can be found in the subfolder results_mature. Here the results for image are saved, e.g. BGH_CY3_01__outline_spots_ 120619.txt. The detection settings are saved in _FQ_batch_settings_MATURE_120619.txt.

16. Averaging of images of individual mRNA molecules

FISH-quant allows averaging the images of individual mRNA molecules to obtain a high quality image. See section 'Averaging of spots' below for more details.

17. mRNA detection in 2D

By default, FQ analyzes 3D images. An analysis of 2D images is also possible, but not supported by all modules. To switch to 2D analysis, go to the menu Settings > 2D/3D detection and select 2d. When you open now a multi-stack image, it will be treated as individual 2D images, and you have to choose which stack you would like to open and process. The other processing steps remain the same. Please note that the batch mode does NOT work for 2D images.

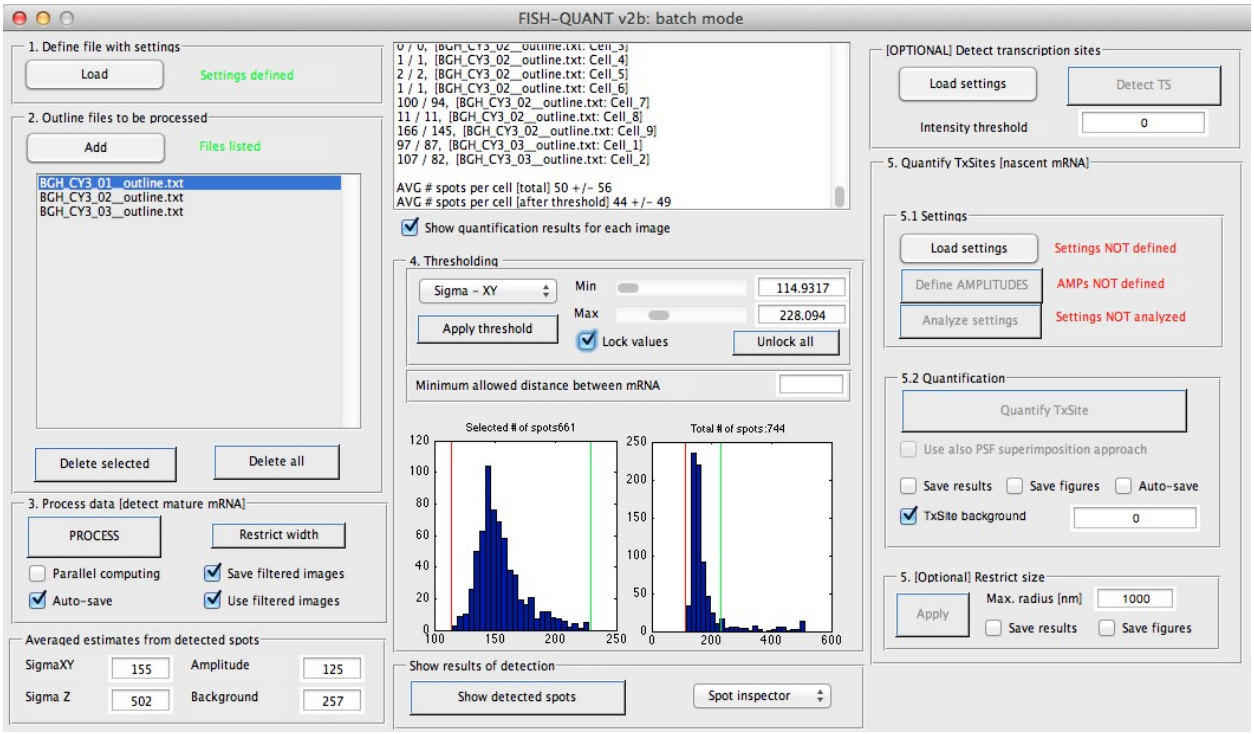
4. Mature mRNA quantification in batch mode

FISH-quant provides the possibility to detection settings for mature mRNA quantification in a batch-processing module to analyze a large number of images automatically with the same settings. It can be started from the menu item *Tools* > *Batch processing* in the main interface of FISH-quant.

An interesting tool to work with larger data-sets is the '**List directory content'**. This allows navigating the different directories and lists the files in these directories. By double-clicking on the files they will be opened in the corresponding tool, e.g. .tif files in the *Outline designer* and spot detection files in the *Spot inspector*. More details can be found in the relevant sections.

TO **STOP A BATCH QUANTIFICATION**, make the main window of Matlab the active window and press CTRL-C. You might have to press this combination a few times until it works. This will produce Matlab error messages but they don't affect FISH-quant.

Restart analysis: to restart an analysis go to the menu item *[FQ] Batch* > *New analysis*.



Using different folders to save results

As in the main interface, tit is also possible to define different folders to store the images, outline, and results (see the corresponding section 'Define folders to save different processing steps' above. These folders are defined independently of the main interface of FISH-quant. This enables working on different folder structures at the same time. Outlines and results will be saved in the currently defined folders. This allows changing these folders during the analysis. This is practical if different analysis strategies are compared, or outlines files are generated during the processing.

The menu item '**Use folders from FQ-MAIN'** will import the folder settings from the main interface. They could then be changed if desired without affecting the settings in the main interface.

The menu item '**Reset folder'** deletes the current folder settings. This can be useful if the folder structure is changed frequently.

Define settings and files that will be processed

First, load the **settings** for mature mRNA detection. Second, the f**iles that will be processed** have to be specified. Usually the **outline of the cells and transcriptions sites** for each image is defined with the outline tool. See 'Outline of cell and transcription site' for more details. Alternatively image files can be added directly and the analysis will be performed on the entire image. It is also possible to use a single outline that will be used for to all images to restrict the analysis to a certain region (*Menu > Advanced > Same outline for all images*). Files can be removed from the list by selecting them (multi-select is possible) and choosing 'Delete selected' or by pressing 'Delete all' all files will be removed from the processing list.

Adding file that will be processed

Next, the outlines files that should be processed can be added. It is also possible to delete a specific file from this list, or to delete all files listed. This is only possible before the actual processing is performed. After that the file-list cannot be changed anymore.

Define how filtered images will be handled

Filtered images are used if the option 'Use filtered image' is enabled. FISH-quant will first check if a filtered image is defined in the outline definition. If not, FISH-quant will check if a filtered image with the default name is present and use it. The default name is the name of image with the suffix '*_filtered_batch.tif',* i.e. for a file called *image.tif* it would be *image_filtered_batch.tif.*) This avoids re-filtering of images and saves processing time. See 'Other options' for how to change this default behavior.

If no filtered image is present or the option 'Use filtered image' is disable FISH-quant will filter the image. By enabling the option 'save filtered images' **the filtered images are saved** with the default name as specified above. FISH-quant saves also a new outline file in a subfolder named _batch within the folder containing the outline files. The name of this file is based on the name in the list of files to be processed, i.e. either the name of the outline file or the name of the image file, and the suffix *_outline_batch.txt* is added.

Parallel computing

Batch-processing uses automatically **parallel computing** (if available) for spot fitting,which helps to reduce the processing time on computers with multiple cores

Autosave

The processing can take some time if a large number of images is processed. To avoid loosing data during unexpected shutdowns the option '**Auto-save**' can be enabled. FISH-quant saves the analysis periodically. This file is saved in the results folder and is named _FQ_analysis_AUTOSAVE_YYMMDD.mat, where the last numbers indicate the date in the format YYMMDD. This file can then be loaded from the file menu (*[FQ] Batch > Load > Analysis results [.mat]*) and the analysis continued with the file where Matlab crashed.

Perform quantification

Batch-processing is **started** by pressing PROCESS. When the processing is finished the data are **thresholded** based on the parameters defined in the detection settings. These settings can, however, be changed. The shown histograms are now for all processed cells. Only LOCKED thresholding parameters will be considered. In the status window the number of all detected spots before and after thresholding are shown. The results **can be inspected** by selecting one file from the list and pressing the button 'Visualization'. The same visualization options as described for the main FISH-quant interface are available.

Saving results

There are a number of different options to save the analysis results.

Results can be **saved** for each processed image **separately.** These files can then be loaded again later. The file-names will start with the name of the image followed by the suffix that can be specified in the interface. For each image a results file as described in Appendix 3 will be saved.

A **summary file for the mature mRNA quantification** can be saved. This file contains the number of transcripts before and after thresholding for each processed cell in a compact format (see Appendix 4). It further contains the area of each cell and (if specified) nucleus.

Example: see __FQ_batch_summary_MATURE_120619.txt in the sub-folder *results_mature.*

It is possible to save **ALL analyzed spots** before or after thresholding in one large file. The user can choose from different ways to label the rows in the summary file: either no label, with the name of the file and the corresponding cell, or with a short file identifier extracted from the file-name. For the latter by default the last four characters of the file-names are used (can be changed in the Advanced menu). This file will be used in the transcription site quantification. When used for this purpose NO label for the rows should be used, otherwise the file cannot be loaded.

Example: __FQ_MATURE_SUMMARY_ALL_spots.txt in the sub-folder *results_nascent.* This file contains the results from 7.5K spots in 150 cells.

Averaging of images of individual mRNA molecules

FISH-quant allows averaging the images of individual mRNA molecules to obtain a high quality image. See section 'Averaging of spots' below for more details.

Loading results of mature mRNA detection

It is possible to load results from previous analysis from the File menu. Results from multiple images can be loaded and re-analyzed.

Other options

- To change the default way to automatically generate the filename of filtered images go to the menu *Settings* > *Filtering*. Here more complex string replacements can be defined. FQ looks for a **UNIQUE** string (a sequence of letters) in the filename of the raw image (the search string) and replaces them with another string to obtain the filename of the filtered image (the replacement string). By default the search string is empty, here the replacement string will be simply added to the filename of the raw image yielding the behavior described above. If the search string is defined, then it will be replaced. This can be useful, for instance, to use deconvolved images as filtering images. The raw image could be could FISH_raw_01.tif and the corresponding deconvolved image FISH_deconv_01.tif. In red a possible search string is indicated, and in green a possible replacement string.

- It is also possible to simply add image files rather than outline files, e.g. TIFF. Here the entire image will be treated as a cell.

- It is possible to use different pre-detection thresholds for different files. The respective (minimum intensity of the spot and minimum quality score) have to be specified in a text file (see Appendix 8). This file can be loaded from the menu *Settings* > *Detection* > *Load threshold for each image*. These detection settings will be used for all specified files; default settings will be used for all other files.

- The experimental parameters (pixel-size, wavelength, …) from the detection setting file will be used rather than the ones save in the individual outline definition files.

5. Averaging of spots

FISH-quant allows averaging the images of the detected spots. Here only spots that pass the thresholding will be considered. This is necessary to obtain a high-quality image of the individual mRNA molecules that can be used for the transcription site quantification.

All functions are available from the Menu 'Average spots' from the main FISH-quant interface and also from the same menu in the batch mode interface. When called from the main interface, the averaging will be performed ONLY on the currently selected cell; when called from the batch interface it will be applied to ALL cells. Stream4.tif

Averaging

When called from the main interface the spots in the currently selected cells will be averaged, when called from the batch mode interface ALL detected spots after thresholding in ALL processed images are averaged. The latter is the method of choice to obtain a really high quality image as needed for the transcription site quantification. When selecting this function a user-dialog opens where several options for this averaging process can be specified.

- The **size of the resulting image.** This is the area around each spot that will be considered in the averaging process. See below for more details.
- During the averaging process the spots are aligned based on the estimated center of the PSF. The center is known with sub-pixel accuracy, the alignment could be done with sub-pixel accuracy as well. This is implemented in FISH-quant as **oversampling (OS)** by sub-dividing the pixels in smaller sub-pixels. The corresponding factors are specified in the fields 'Factor OS'. When using an oversampling factor of 3 each pixel is sub-divided in 3*3*3 = 27 sub-pixels (see Appendix 10). By default, however, no oversampling is used.
- A **local background subtraction** can be applied. This background is obtained from the fit with the 3D Gaussian and subtracted from the image prior to the averaging.

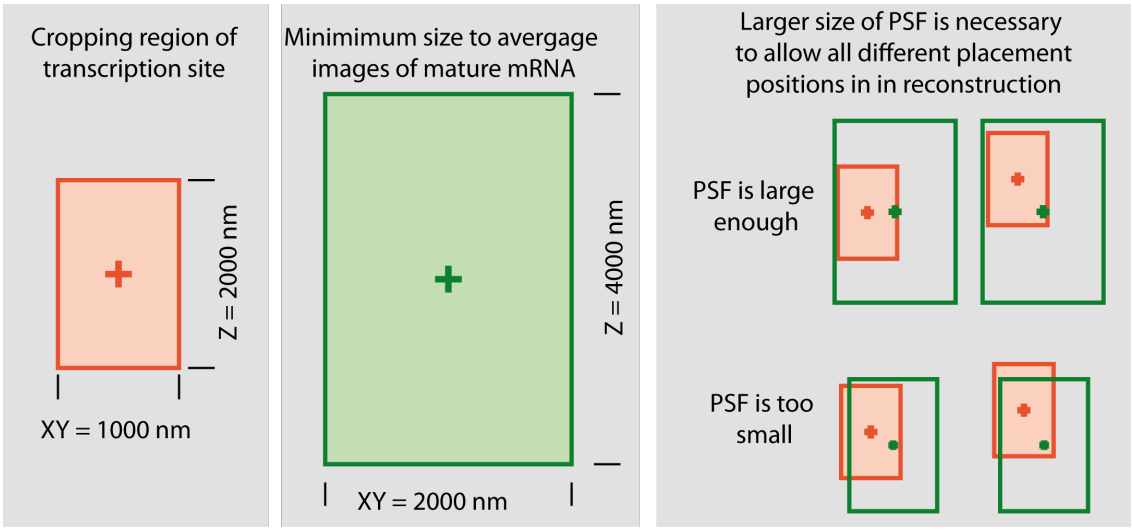
Size of averaging region

The averaged image of the mRNA molecule is used for the different transcription site quantifications. In the quantification approach based on the superimposition of PSFs an image of the transcription site is generated by iterative placement of this averaged image until the best possible match with the actual image of the transcription site is achieved. FISH-quant crops the image of the transcription site automatically to perform this analysis on a smaller sub-region. The averaged image of the mRNA has to be a least twice as large as this region to guarantee that is large enough to be placed anywhere in the reconstruction (see cartoon below for illustration). Otherwise problems could occur when the mRNA is placed towards the edge of the reconstruction (see plot below).

By default the transcription size is cropped from the brightest pixel +/- 500 nm in XY and +/- 1000 nm in Z. This means that the averaging region has to be at least +/- 1000 nm in XY and +/- 2000 nm in Z. A larger averaging region in XY has the additional advantage that the background of the averaged image can be determined with higher confidence (see section 'Define BGD of PSF' in the chapter about transcription site quantification).

Spots are only considered in the averaging process when the entire sub-region can be extracted. If too many spots are not considered either a smaller averaging area has to be specified or larger images have

to be collected. If a smaller averaging area is defined, the size of the cropping region has to be adjusted accordingly.



Saving/viewing

All images can be shown in ImageJ or directly saved as a tiff file from the respective sub-menus. By default the image will be save to the results sub-folder (and not the image sub-folder).

Fitting

The resulting spot can be fit with the 3D Gaussian function. The size of the fitting region is the same as used for the pre-detection of spots.

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.

6. Detection of transcription sites in batch processing

It is possible to detect transcription site in the batch-processing module. First, the detection settings have to be defined and saved in the outline-module (see corresponding section). These settings can then be loaded into the batch-processing module with the button ‘ Load settings’ in the corresponding panel. These settings can then be applied by pressing ‘Detect TS’. The currently defined minimum intensity of the transcription site will be displayed in the text box ‘Intensity threshold’. This is the

FISH-quant will generate new outline files in the specified outline-folder. If you want to test different intensity thresholds you can specify a new folder for each threshold and inspect the results afterwards.

Saving/viewing

FISH-quant provides simple browser (see section ‘Inspection of directory content’) that allows inspecting outline files in a directory. This tool can be used to quickly inspect the automatically generated outlines and change regions that were assigned erroneously.

7. Quantification of individual transcription sites

The quantification parameters can be defined and tested in this GUI and then be saved in a text file and used in the batch mode to process a large number of cells. In the following section these parameters are explained and also examples based on the example data set are given.

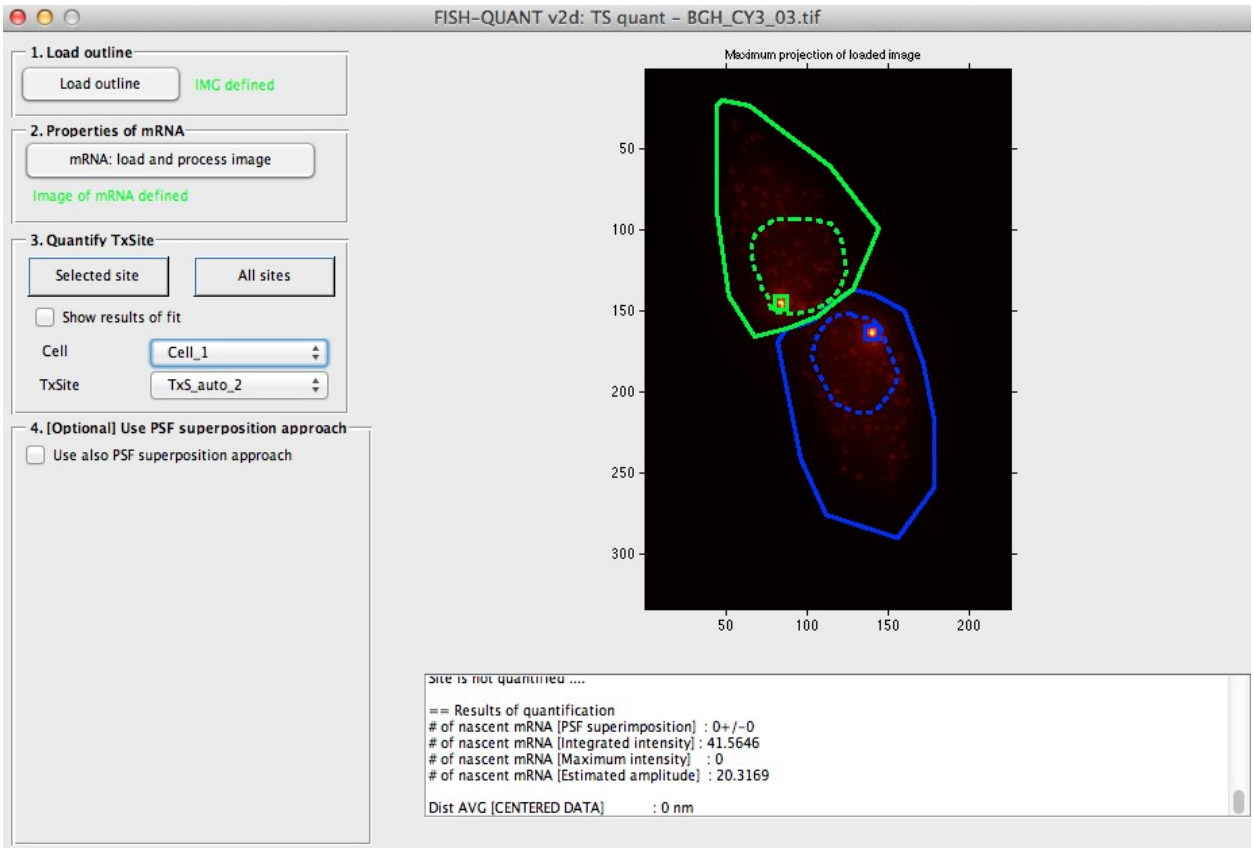
Different quantification methods

FISH-quant quantifies the amount of nascent transcript at transcription sites with four different methods

1. **Comparison of integrated intensity** of transcription site and individual mRNA molecules
2. **Superimposition of PSFs** to reconstruct an image of the TS
3. Comparison of **estimated amplitude** of transcription site and individual mRNA molecules
4. Comparison of **maximum intensity** of transcription site and individual mRNA molecules

In simulations we found that the method based integrated intensity and the PSF superimposition approach yield accurate estimates also for spatially extended TS. The method based on the **integrated intensities** is computationally faster and works well for sites that are larger than individual transcripts but still relatively homogenous in shape. We generally recommend using this option.

Additionally, the PSF superposition approach can be used and the results compared as an internal quality control. The other two methods lead to an underestimation of the number of nascent transcripts for large and spatially extended transcription site.



Quantification with integrated intensity

Here only the averaged image of all mRNA has to be specified.

Button 'Image of mRNA'

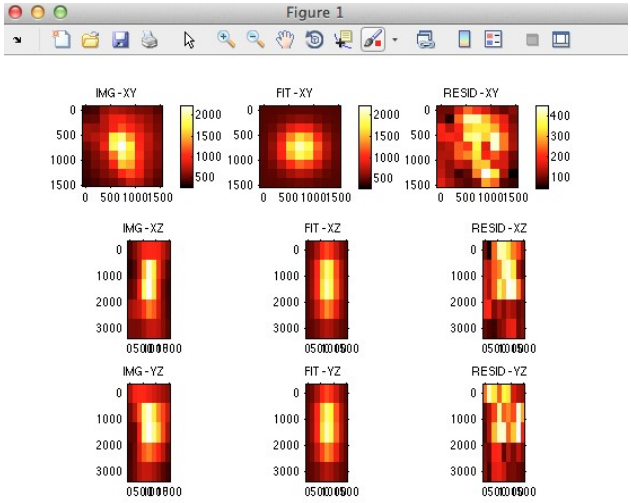
With this button the TIF file with the averaged image of an individual mRNA is loaded. The section 'Averaging of spots' explains in details how to obtain such an image.

Example: the averaged image of the PSF in the example data set (_mRNA_AVG_ns.tif)

We further recommend to **check that the region that's automatically cropped around the TS is appropriate**. This can be changed in the menu [FQ] TxSite > Option. When the region is changed, the image of the RNA has to be loaded again. For the definition of the size of the region a balance has to befound. The regions should be

- **Large enough** to encompass even the brightest TS but also
- **Small enough** that the fit of dim TS with the 3D Gaussian still works. If the cropping is too large, the fit with the Gaussian can fail.

This can be checked by enabling the option '*Show results of fit'*. This will show several figures summarizing the different parts of the quantification. The first figure shows 3 columns. The first column is an image of the TS with maximum intensity projections along the three axis. The second column the results of the fits with the Gaussian. The third column the residuals, e.g. the difference between the data and the fit. The region in the example is large enough that the entire signal is capture (background around site is included) but not so large that background gets dominant. Fit also nicely captures the TS.



Quantification

Then the quantification can be started for either the currently selected transcription site or all sites in the image by pressing the respective button.

Loading and saving analysis setting

The analysis settings can then be save from the main menu (see Appendix 5). These settings can then be loaded again and modified. They are also used in the batch-processing module. In this file the path-name for the averaged image of mRNA and the results file are specified. When settings are loaded FISH-quant first checks if the files can be found in this folder. If not (or no folder is specified) FISH-quant will check in the same directory that the settings file. If it can't locate the files there the corresponding files

have to be specified again. To avoid this problem it is advised that these files are stored in the same folder than the settings file.

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

Saving results of quantification

The results of the transcription site quantification can only be saved if the quantification was performed with the option to quantify all sites at once. Only then the corresponding Menu will be enabled. The size restriction for the PSF superimposition approach will be considered only if it has been applied to all sites.

***Optional*: quantification of background signal**

FISH-quant allows the automated detection of transcription sites, e.g. with LacI. In this case also sites that are not actively transcribing can be detected and will be analyzed. The FISH image will therefore contain no signal and only background. This can, however, lead, to problems with the Gaussian fitting routine. We found that such background regions are often fit with very large Gaussians with a small amplitude. If these fitting results are used to calculated the integrated intensity and infer the number of nascent transcripts erroneously large estimates are obtained. FISH-quant therefore attempts to detect these wrong fits based on two criteria. If one of them is met, the estimated number of nascent transcript is set to 0.

1. **Size only**. Fit will not be considered if the estimated σ of the Gaussian (in XY and Z) are 3-times larger than the estimated σ for the individual mRNA molecules.

2. **Size and Shape**. Fit will not be considered if the estimated σ of the Gaussian (in XY) is 3-times larger than the estimated σ for the individual mRNA molecules AND if the estimated amplitude is smaller than twice the amplitude of the mRNA.

If outlines are processed that contain transcription sites with no signal, we quantify some of these sites manually to verify that these criterion work well. If necessary, they could be adjusted in the function `FQ_TS_analyze_results_v8`.

Optional: quantification with PSF superposition approach

After enabling this option new panels needed for the defintion of additional parameters are shown.

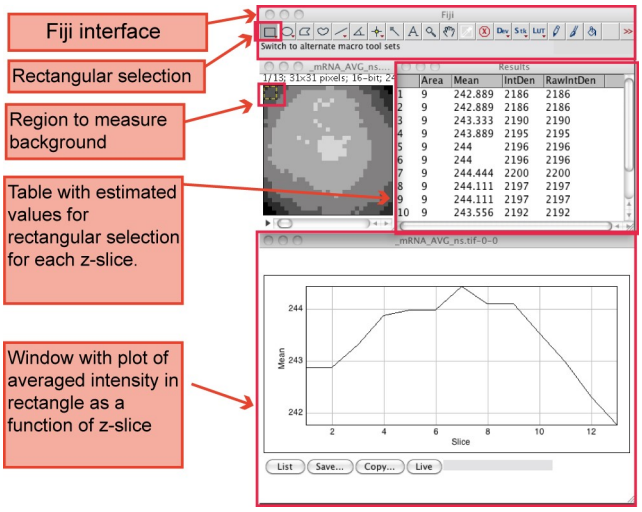
Button 'Define BGD'

Here the background intensity of averaged mRNA can be specified. This is **ONLY** important for the PSF superimposition approach. If this approach is not used the background value can be set to an arbitrary value since the fitting to estimated the integrated intensity will be done on the raw image.

The background can be defined either as a scalar value or as an image. If a **scalar value** is defined, this value will be subtracted from each pixel. If an **image** is defined, this image file has to have the same dimensions as the image of the mRNA. The background subtraction will then be performed pixel-by-pixel.

We usually use a scalar value that we measured in regions of averaged image of the mRNA without signal away from the center. **This can be measured in Fiji** (http://fiji.sc/) in a few steps.

1. Open averaged image of mRNA in Fiji
2. Use rectangular selection (first item on the left under the menu).
3. Draw a small region in one of the corners, e.g. 3-by-3 pixels.
4. Got to Image > Stacks > Plot Z-axis Profile. This will calculate the averaged intensity within this region for each z-slice. Fiji will display two new windows. One with a table summarizing the results, one with a plot showing the averaged intensity for each plane. In our example this value is between 242 and 244, so we choose 243.



It is better to define the background slightly too larger than too small. When using a too small background the background subtracted image will still have some residual intensity in background regions. This will affect the reconstruction and lead to an underestimation of the number of nascent transcripts. With a slightly too large background some mRNA's might be placed outside of the actual transcription site to give a better description of the background. These can then be removed by confining the size of the transcription site (see below).

Example: we use a scalar value of 243 for the background subtraction and restrict the size of the transcription site to 1000nm.

Button 'Amplitudes'

Here the fitting results of the individual mRNA are loaded either from the current analysis or from a file containing the summary of the spot detection (can be saved in the batch mode for all detected spots). FISH-quant extracts the amplitudes of the fits and the maximum intensity of the individual mRNA molecules. These values are used in the simpler quantification when they are compared to the values obtained for the transcription site. The extracted amplitudes are used in the PSF superimposition approach to determine the brightness of the individual images that will be used in the reconstruction.

FISH-quant can consider either the distribution of the estimated amplitudes or simply their mean value. This can be specified in a question dialog that will be displayed. If TxSites are analyzed only with the integrated intensity, the mean value is sufficient. If TxSites are analyzed with the superposition approach and a large number of mRNAs were analyzed we recommend using the distribution, otherwise the mean value.

For the quantification with the integrated intensity only the averaged image of the mRNA is necessary. So if no results are specified, only this quantification method can be used.

Example: file __FQ_MATURE_SUMMARY_ALL_spots.txt. It contains the detection results of 7.5K individual mRNA molecules.

Button 'Analyze settings'

Then the settings for the mRNA image are processed and prepared for the quantification. As a quality control the image will be fit with the 3D Gaussian function and the results displayed. FISH-quant will then simulate images of mRNA based on the distribution of amplitudes and – if specified - the different sub-pixel placements. Each simulated image will be fit with the 3D Gaussian with the same cropping range used to analyze the transcription sites. From these fits the averaged value of the amplitudes and the sigma's are calculated. They are then used to estimate the integrated intensity. See the command line for status of this process.

[Optional] Define BACKGROUND of the transcription site

The quantification with the PSF superimposition approach assumes a constant background value for the transcription site. Two options are available to determine this background

- By default the background value can be **determined automatically** by disabling the option *TxSite background*. The background will be then determined individually for each cell. First, the outline of the cell (or if present the nucleus) is used to extract all pixels in the cell that are in the same z-slices than the transcription site. Then the median and the standard deviation of all pixel intensities are calculated. With these estimates a range of background values is determined and tested in the PSF superimposition approach. By default 10 values cover the *[median – 3*stdev, median + 1*stdev]*. These settings work well for cells with homogenous background, for other cells the settings have to be adjusted accordingly (see *Quantification options* below). We use the automatic options to determine the background values. By using a smaller range or less values the quantification can be restricted a background value close to the estimated median value.

- The user can define **a certain range of background values** and FISH-quant tests the quantification for each of them and chooses the value that gives the best reconstruction, i.e. with the smallest squared sum of residuals. For example type '260, 270, 280' (without the ') to test these three intensity values and enabling the option *TxSite background*.

A good way to check if the **settings are ok** is to enable the 'Show results' option. One of the plots summarizes the reconstruction for the different tested background values. One curve shows the residuals of the reconstruction as a function of the different background values. This curve should have an U-shape with the minimum indicating the best background value. This value should also be close to the background of the cell in the vicinity of the transcription site.

Restriction of size of the transcription site

This post-processing step allows setting an upper limit to the size of the transcription site for the PSF superimposition approach. It can be applied either to the currently selected transcription site or to all sites if they have already been processed. FISH-quant calculates the center of mass for all mRNA molecules in a transcription site. Then the distance of the molecules to this center is determined and molecules excluded if this distance exceeds the defined upper limit. This option is useful when individual molecules are placed outside of the transcription site. In plot [6], excluded molecules are indicated by blue crosses.

QUANTIFICATION

Again, the quantification can be started for either the currently selected transcription site or all sites in the image by pressing the respective button. The quantification based on the PSF superimposition approach can be enabled with the corresponding enabling button. If the Parallel computing toolbox is installed the '**Parallel computing'** option helps to speed up the computation on machines with multiple cores.

When the option '***Show results'*** is activated a number of plots are shown during the quantification:

- [1]: Figure that shows the transcription site after cropping to judge if the cropping area surrounding the transcription site has the appropriate size.
- [2]: This plots is shown when multiple background values are tested. The first panel shows the best residuals for each background value. Ideally this curves has a U-shape and the minimum indicates the best background value. The second panel shows the number of placed mRNA's, i.e. the number of nascent mRNAs, for the different background values. The third panel shows an overlay of the first two plots.
- [3]: Various histograms showing the used amplitudes for the reconstruction with the PSF superimposition approach.
- [4]: A plot showing the residuals for the reconstructions. The minimum of these plots is then interpreted as the number of nascent mRNAs.
- [5]: A histogram of the residuals with only a background image and with the best reconstruction. The residuals of the best reconstruction should be centered around zero.
- [6]: Images of transcription sites, the best reconstruction, the absolute residuals, and the positive and negative residuals. Green crosses indicate the location of the placed mRNAs.

Inspection of results

Results for analyzed sites can shown by enabling the 'Show results' and selecting the site from the list. Plots 4, 5, and 6 from above are shown as well a summary of the quantification in the Matlab workspace and the status interface of FISH-quant.

Visualization of results in ImageJ

The results of the quantification can be visualized in ImageJ by first selecting the site and then pressing 'Visualize (ImageJ)'. Two images are then generated. In the first image the Z-stack of the transcription site is shown next to the z-stack of the best reconstruction. In the second image the residuals are shown: positive residuals are shown in red, negative residuals in green.

Note: a useful function in ImageJ is reslice (*Image* > *Stacks* > *Reslice [\]* …). This allows flipping the stacks, i.e. rather than having different xy slices, xz slices are shown.

Quantification Options

For the PSF superimposition approach a number of default settings are defined which can be modified.

Placement

Determines the strategy to place the individual mRNAs molecules. After each round the reconstruction is subtracted from the transcription site to obtain the residuals. They mRNA is then either placed at the location of the maximum intensity of the resulting image (2) or a weighted random positions (1). In the latter each voxel has a probability to be selected for the next placement that is proportional to its squared residuals, i.e. it is more likely to place a Gaussian at a location with high than low residuals.

Residuals

Used residuals in PSF superimposition approach: Squared sum of residuals (1), absolute sum of residuals (2).

Number of runs

Number of individual reconstruction runs.

Number of preliminary runs

Number of runs to determine the maximum number of placed Gaussians and - if necessary - the best background intensity.

[BGD auto]
Parameters to determine the automatic background estimation. For meaning of parameters see section 'Define BACKGROUND of the transcription site' above.

[CROP] in XY +/- nm and *[CROP] in Z +/- nm*
Determines size of region around the identified center of the transcription site that will be considered in the analysis. This size should be adjusted to the typical size of the transcriptions site in the data. It should contain the entire transcription site without too much background signal. We recommend verifying the size for a small and large sites.

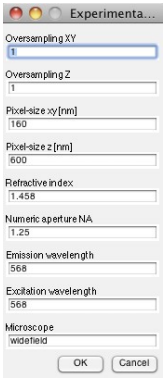
[AUTO-DETECT]
Settings for the auto-detection of the transcription site. See section 'Outline of cells and transcription sites' for more details.

[Size of region to sum intensity]
Determines size of region around the identified center of the transcription site that will be used to sum up pixel intensities. This is currently not used for any analysis and is intended for a future release.

Menu: Parameters

Here various parameters describing the averaged image of the mRNA are specified: oversampling factor in XY and Z (default 1), the pixel-size in XY and Z in nanometer, the refractive index, the NA of the objective, the emission and excitation wavelength and the type of the microscope.

Example: figure on the right shows the parameters as set for our analysis.



8. Quantification of transcription sites in batch mode

TO **STOP A BATCH QUANTIFICATION**, make the main window of Matlab the active window and press CTRL-C. You might have to press this a few times until it works. This will produce Matlab error messages but they don't affect FISH-quant.

Basic workflow

In the GUI of the batch mode the **settings file for mature detection nd the outline definitions** have to be loaded first (see 'Mature mRNA quantification in batch mode'). Loading the outline files is sufficient; it is NOT necessary to process these files. Alternatively it is possible to load the results file of the mature mRNA quantification. This is useful if a summary file will be generated that contains for each cell the amount of mature and nascent mRNA (see below). Next the **settings for the transcription site quantification** have to be specified. In this file the averaged image of the mRNA and the summary file of the mature mRNA detection are specified. Either file can still be changed. Please note that if FISH-quant can't locate these files the user has to specify them again. This could happen if the indicated path in the settings file is not valid. Lastly these settings have to be analyzed (see section about transcription site quantification for more details).

Example: load the settings file for mature detection *_FQ_settings_MATURE.txt* from the subfolder *results_mature*. Load the results file of the mature mRNA quantification from the subfolder *results_mature*. Load the settings file for transcription site quantification *_FQ_settings_NASCENT.txt* from the subfolder *results_nascent*.

Save summary of quantification

Enabling the options 'Save results' and 'Save figures' allows saving more detailed information about the quantification. FISH-quant will generate a folder called TS_QUANT_YYMMDDD in the results folder, where the last string is the current date. When the option '**Save figures'** the same images as described above in the section 'Visualization of results in ImageJ' are saved for each analyzed transcription site. Please note that this will increase the overall run time considerably. When this option is enabled Matlab will open and close windows during the reconstruction process. Please don't close these windows manually since this will cause error messages. When the option '**Save results'** is enabled a text file will be generated where the most important parameters of the transcription site reconstruction are saved. In addition a new outline file is save if the auto detection option was activated.

Autosave

The transcription site quantification with the PSF superimposition approach can take several hours. To avoid loosing data during unexpected shutdowns the option '**Auto-save**' can be enabled. Matlab saves the analysis periodically (see 'Loading and saving analysis' below). This file is saved in the results folder and is named _FQ_analysis_AUTOSAVE_YYMMDD.mat, where the last numbers indicate the date in the format YYMMDD. This file can then be loaded and the analysis continued where Matlab crashed.

Start quantificaiton

Then the quantification for all transcription sites can be started.

Restriction of size of transcription site

The size of the transcription sites can be restricted for the PSF superimposition approach as described in the section about transcription site quantification. New status files and figures files can be saved for this restriction by enabling the corresponding options. The same file as described above will be saved in a

subfolder called TS_QUANT_REST_1000nm__ YYMMDD, where the value after REST indicates the maximum size used for the restriction.

Note: this can take a few minutes for a larger number of files when new status files and figures are saved as well. When new figures are generates Matlab figures will open and close windows. Please don't close these windows manually since this will cause error messages.

Loading and saving analysis

The current status of an analysis can be saved and loaded from the menu *[FQ] Batch> Load/Save > Analysis results [.mat file]*. The entire analysis is saved in a Matlab .mat file that can be later loaded again and the analysis can be continued.

The same file is save when the option '**autosave'**. Here FISH-quant will save the current status of the analysis after each processed image. You can load this file and continue the analysis where it stopped in case the computer crashes.

Note: if this file is loaded on a different computer the corresponding path settings have to be updated.

Saving results

Quantification results can be saved in different summary files (Menu *[FQ-Batch] > Save)*

Summary: nascent mRNA
Gives a detailed summary for each transcription site (Appendix 6). 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 (Appendix 7). 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.

Autodetectio of transcription sites

It is also possible to only perform the **automated detection of the transcription sites**. For each image a new outline file will be generated and save with the same name as the original file with the addition suffix '_TS_auto_550', where the number specifies the intensity threshold used for the detection. The files are saved in a subfolder called *_TS_detect* within the folder containing the original outline files. These files can then be inspected and processed with the quantification. It is not necessary to load the quantification settings for the transcription site to perform the auto detection. Simply define the intensity threshold and enable the checkmark to be able to use the function.

9. Advanced features

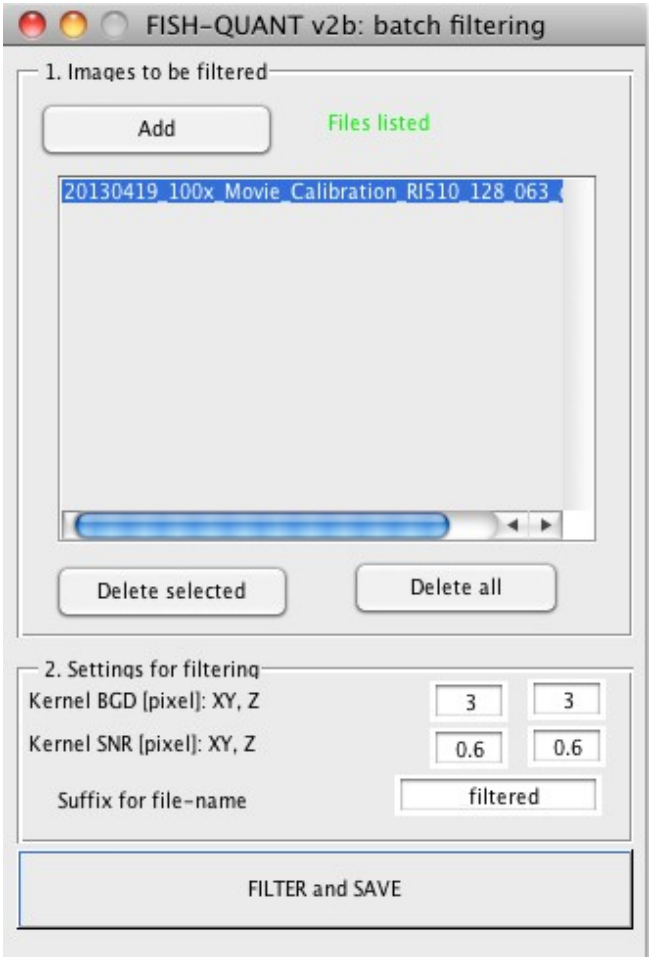
Batch filtering

FISH-quant provides the possibility to filter a large number of images in a batch mode. This can be useful to process files that will be used for the automated detection of transcription sites, e.g. images with LacI. The corresponding GUI (shown below) can be called in the main interface from the Tools menu and 'Batch filtering'.

FISH-quant will automatically import the filtering settings from the main interface (see corresponding section above). It will also use the same experimental parameters and use the specified folder for images.

The user first specifies the images that should be filtered. The filtering settings can be changed if needed. Each filtered image will be automatically saved with a new file name. This file name is generated from the name of the image and an added suffix (by default _filtered).

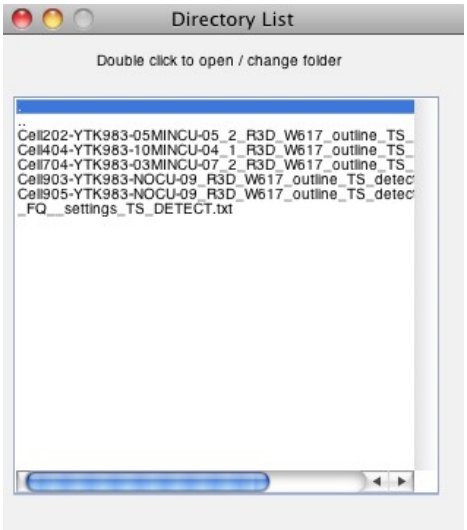
Press **FILTER and SAVE** to process all specified files.



Inspection of directory content

FISH-quant provides a tool to inspect outline files from a directory browser. This tool can be opened from the main interface menu *Tools* > *List directory content*. **Before** opening this tool the folder definitions (images, outlines, results) have to be set.

This will bring up a small interface as shown on the right. You can navigate through the folders by either opening folders or pressing '..' to move up one directory.



You can double-click on the following types which will then be automatically be opened in the corresponding FISH-quant module. This GUI is also resizable.

- **Image files (TIF).** Will be opened in the Outline designer and used as FISH images. Convenient to design outlines of many images files.
- FISH-quant **outline files** (must contain 'outline' in the file-name). Will be opened in the Outline designer. This is convenient to inspect automatically generated outlines files, e.g. for detection of nuclei or transcription sites. If the detection was not good, the outlines can then be directly changed and saved.
- FISH-quant **results files** (must contain 'spots' in the file-name). Will be opened in the Spot inspector. Convenient to inspect the results of a batch analysis.

Automated outline generation based on DAPI signal

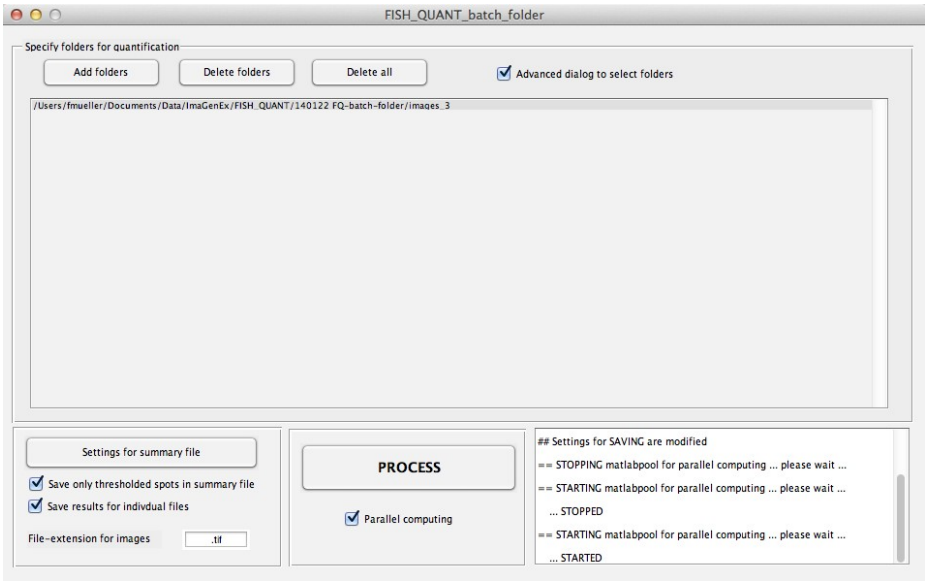
It is possible to automatically generate cell outlines based on the DAPI signal (script '*SCR_outline_auto_DAPI_v2*' in scripts subfolder). This approach can be used when mRNA is predominately located in the nucleus or only the molecules in the nucleus are of interest. Here FQ will first detect the nuclei based on the DAPI signal (see section 'Outline of cells and transcription sites' for more details). The detected outlines are then used as the cellular outlines. More details can be found in the annotated script.

Batch processing of multiple folders

This GUI allows specifying multiple folders that will be processed automatically (Tools > Batch processing – folders). Each folder has to contain

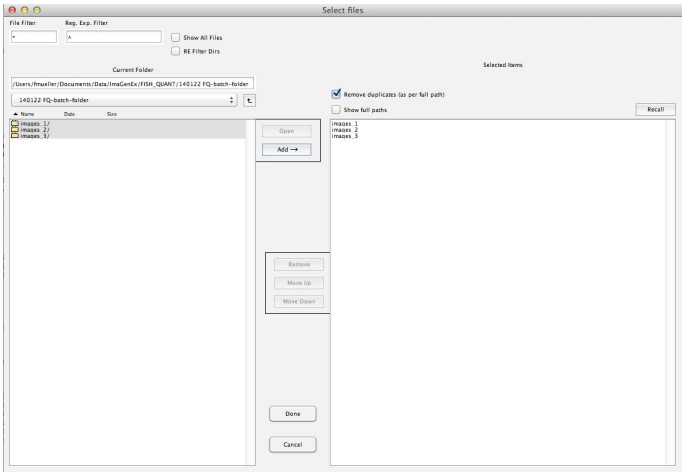
1. A settings file for the mature mRNA quantification (file-name has to contain the string *_settings_*). This allows to have different settings files for different folders. However, each folder can contain only ONE settings file.

2. Image files. Images are identified based on their file-extension which can be changed in the GUI (default is .tif).



The user has to first specify the different folders that will be processed. Folders can be added or removed with the corresponding buttons. To add files two different possibilities exist. If the option 'advanced dialog …' is selected, the dialog shown below will be used to select folders.

Here folder can be selected on the left and added to the selection with the arrow. Multiple folders can be selected, also with different parental folders. Press DONE when selection is finished.



If this option is disabled, then one folder at a time can be selected.

After the folders are selected, a few options can be specified

- FISH-quant will save for each folder a **summary file with the parameters of all detected spots** in all images. Here the user can choose if only the spots after thresholding are saved, or all spots. Thresholding parameters have to be defined in the settings file. File is saved in the same folder as the images.

- In the summary file each row is labeled with a **unique identifier** for the underlying image. This identifier is extracted from the file-name. By default the last 4 characters of the name (without the

extension are used). This can be changed with the dialog that opens when Pressing '*Settings for saving summary file'*).

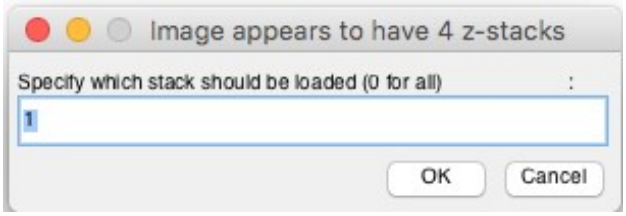
- It is also possible to save the **results file for each of the images**. By default this option is disabled. Files are saved in the same folder as the images.
- **Parallel computing** can be used to speed-up the Gaussian fitting.

Press PROCESS to start the processing.

10. Troubleshooting

Loading images: warning message “Image appears to have … z-stacks”

This message can occur when loading an image in the presence of a Metafile (e.g. files with the same name as the image but with different extensions, e.g. nd on Nikon microscopes). Deleting these metafiles or moving them to a different folder usually resolves the problem.



Error when opening images images: “found muliple channels …”

See error above. Deleting or moving the metafiles often resolves this problem.

As an alternative, you can also change the function that opens images. In the Menu *Settings>Function to open images*. Here, select “tiffread” as an alternative. This function is a little slower but usually allows solving this problem.

Opening of images: error message or running out of memory

This can be caused by the default routine used by FISH-quant to open images (bfopen). This default behavior can be changed from the Menu *Settings>Function to open images*. Here you can select “tiffread” as an alternative. This function is a little slower but usually allows solving this problem.

Windows: Folder with FISH-quant source code is encrypted

FISH-quant is provided as a ZIP archive generated in MacOS. On rare occasions problems might occur when unzipping this archive under Windows 7 with the built-in unzipper. Here the folders are **encrypted** and are shown in green. This can be avoided by using the free program 7-Zip (http://www.7-zip.org/).

Working with network folders under windows

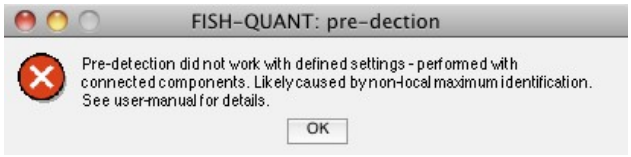
Often images are stored on network folders. When FISH-quant is running on windows machines these folders are sometimes not visible when defining the folder structure or loading image. We found that the following little trick can help. First, go to windows explorer and open the corresponding network folder. Copy the full name of this folder. Second, in FISH-quant define the root folder and past the copied name from the first step. FISH-quant then recognizes this folder and the other folders (images, outlines, ...) can then be defined directly.

Filtering takes a very long time

This filtering is performed on the entire image independently of the defined cells. This step can take a long time for large stacks. Cropping the images leads to a faster filtering. Filtered images can also be saved and then an updated outline-file can be saved that will reference to this save file so the filtering step doesn’t have to be performed again.

Error message when performing pre-detection

The error message shown on the right can be displayed when performing the pre-detection.



This error is caused when the Matlab routines necessary to perform the pre-detection with the **non-maximal suppression method** cannot be executed. FISH-quant will perform the pre-detection with the **connected components method** instead. Two possible errors are:

1. The actual Matlab functions performing the local maximum detection (parts of the Piotr's Toolbox[1]) are not properly compiled. To compile them, open Matlab and type in the command window 'toolboxCompile' (without the apostrophe) and hit enter. This usually takes a few minutes and has to be performed only once.

2. FISH-quant is running on older versions of Matlab. Here only an upgrade to a newer version of Matlab solves the problem.

11. Appendix

Appendix 1 – definition of outline files

These files are defined for each recorded stack. They define the parameters of the experiment and for each cell in the field of view the outline of this cell and the location of the transcription site(s). **They have to be stored in the same directory than the image**. An example for such a file can be found below (grey box).

ROW	FILE
1	FISH-quant v2d
2	File-version 3D_v1
3	OUTLINE DEFINITION, 19-Sept-2014
4	COMMENT Outline definition performed in FISH-quant (Main program)
5	IMG_Raw BGH_CY3_02.tif
6	IMG_Filtered
7	IMG_DAPI BGH_DAPI_02.tif
8	IMG_TS_label
9	FILE_settings FISH_test_data__settings.txt
10	PARAMETERS
11	Pix-XY Pix-Z RI Ex Em NA Type
12	160 300 1.458 568 568 1.25 widefield
13	CELL_START Cell_1
14	X_POS 117 98 85 100 115 138 160 202 236 235 218
	188 165 150 131 122
15	Y_POS 215 247 283 311 331 354 373 369 308 253 211
	193 194 199 204 209
16	Z_POS
17	CELL_END
18	Nucleus_START Nucleus_manual
19	X_POS 160 147 146 143 141 141 142 146 151 160 170
	183 192 198 202 202 202 197 188 182 166
20	Y_POS 244 252 260 270 286 301 312 329 336 344 344
	342 335 325 306 292 276 264 250 243 241
22	Z_POS
22	Nucleus_END
23	TxSite_START TxS_auto_1
24	X_POS 172 180 180 172
25	Y_POS 250 250 258 258
26	Z_POS
27	TxSite_END

- **Line 1-3** Header row, specifying version number, file, and date of analyis
- **Line 4** is reserved for a user-comment and starts with the key word **COMMENT**
- **Line 5** contains the name of the raw image and starts with they key word **IMG_Raw**
- **Line 6** contains the name of the filtered image and starts with they key word **IMG_Filtered**. Empty if no imaged is defined.
- **Line 7** contains the name of the DAP image (or any other nuclear stain) and starts with they key word **IMG_DAP.** Empty if no imaged is defined.
- **Line 8** contains the name of the image with the independent label of the transcription site and starts with they key word **IMG_TS_label.** Empty if no imaged is defined.
- **Line 9** contains the name of the file with the settings for the mature mRNA detection and starts with they key word **FILE_settings.** Empty if no settings file was saved.
- **Line 10** is the header to indicate that the experimental parameters will come next

- **Line 11** contains a description of the parameters in line 11.

Pix-XY	Pixel-size in XY in nanometer
Pix-Z	Pixel-size in Z in nanometer
RI	Refractive index of medium
Ex	Excitation wavelength of fluorophore
Em	Emission wavelength of fluorophore
Type	Type of microscope. Currently only widefield is supported.

- **Line 12** specifies the experimental parameters.
- **Line 13** specifies the first cell. Starts with keyword **CELL_START**, followed by a unique identifier of this cell.
- **Lines 14-16** specifies the coordinates of the polygon describing the outline in X,Y,Z (reserved for future development)
- **Line 17** specifies the end of the cell definition with the key-word CELL_END
- **Lines 18-22** specify the polygon of the nucleus for the above-defined cell**.** Can be omitted if no nucleus is defined. ONLY one nucleus can be defined per cell.
- **Lines 23-27** specify the polygon of the transcription site (TS) for the above-defined cell. Can be omitted if no TS is defined, or repeated if multiple TS are defined.

Notes

- It is not necessary to defined nuclei or transcription-sites. These blocks can be omitted.
- Per cell only one nucleus can be defined.
- Multiple transcription sites can be defined per cell by repeating the block from line 18-20 and changing the identifier in line 18.
- Multiple cells can be defined by repeating the block from line 12 to line 20 with unique cell label.
- Files are saved as .txt files; All values are **separated by tabs**; no empty rows are allowed between definitions.

Appendix 2 – definition of files summarizing results of spot detection

These files are defined for each cell in the recorded stack. Relevant information from outline definition files is repeated (experimental parameters, outline of cell and transcription site. An example for such a file can be found below (grey box).

ROW	FILE											
1-27	See Appendix 1											
28	SPOTS_START											
29	Pos_Y	Pos_X	Pos_Z	AMP	BGD	RES	SigmaX	SigmaY	SigmaZ	Cent_Y	Cent_X	Cent_Z
		MuY	MuX	MuZ	ITERY_det	Y_det	X_det	Z_det	Y_min	Y_max	X_min	X_max
		Z_min	Z_max	INT_raw	INT_filt	SC_det	SC_det_norm		TH_det	TH_fit		
30	19458.4	20127.7	5832.73	180.341	476.227	25197.6	181.736	181.736	882.762	316.289	318.481	1203.91
		258.367	287.675	1332.73	10	123	127	20	121	125	125	129
		16	24	581	91	20.2438	0.0732658	1	1			
31	12946.3	11145.3	6434.95	70.446	485.745	35573.8	910.826	910.826	965.181	320.5	319.634	1204.4
		466.284	265.328	1334.95	11	81	71	22	79	83	69	73
		18	26	560	33	7.79373	0.0282069	1	1			
32	SPOTS_END											

- **Line 1-27** Same information as in outline file. So results of spot detection contain all information about the cells and can serve as outline files as well.

- **Line 28** Key word SPOTS indicates that results of spot detection start
- **Line 29** Header for spot detection

Pos_Y	Estimated position of center in image (Y, in nanometer)
Pos_X	Estimated position of center in image (X, in nanometer)
Pos_Z	Estimated position of center in image (Z, in nanometer)
AMP	Estimated amplitude
BGD	Estimated background
RES	Squared sum of residuals
SigmaX	Estimated width = sigma of Gaussian (X, in nanometer)
SigmaY	Estimated width = sigma of Gaussian (Y, in nanometer)
SigmaZ	Estimated width = sigma of Gaussian (Z, in nanometer)
Cent_Y	Centroid of sub-image (Y, in nanometer)
Cent_X	Centroid of sub-image (X, in nanometer)
Cent_Z	Centroid of sub-image (Z, in nanometer)
MuY	Position in sub-image (Y, in nanometer)
MuX	Position in sub-image (Y, in nanometer)
MuZ	Position in sub-image (Y, in nanometer)
ITER	Iterations used for fit
Y_det	Pre-detected position in image (Y, in pixel)
X_det	Pre-detected position in image (X, in pixel)
Z_det	Pre-detected position in image (Z, in pixel)
SC_det	Pre-detection score
SC_det_norm	Normalized pre-detection score (to maximum)
Y_min	Start of sub-region for detection in image (Y, in pixel)
Y_max	End of sub-region for detection in image (Y, in pixel)
X_min	Start of sub-region for detection in image (X, in pixel)
X_max	End of sub-region for detection in image (X, in pixel)
Z_min	Start of sub-region for detection in image (Z, in pixel)
Z_max	End of sub-region for detection in image (Z, in pixel)
INT_raw	Intensity of pre-detected spot in raw image
INT_filt	Intensity of pre-detected spot in filtered image
TH_det	Thresholded after detection (1 = good, 0 = bad)
TH_fit	Thresholded after detection (1 = good, 0 = bad, -1 manually removed)

- **Line 30- 31** Results of spot detection. Each line corresponds to one spot (2 in this case)
- **Line 32** Keyword **SPOTS_END** to indicate end of spot detection results for this cell.

This block defining the detected spots is placed immediately after the outline definition of each cell. if multiple cells are defined, this block will be placed after each one specifying the spots in the respective cell.

Appendix 3 – file with settings for mature mRNA detection

This files defines all parameters describing the processing and detection process. It can be read-in by the batch processing tool. Settings are stored in a simple format. Each property starts with the name followed by a '='and the actual value. There is NO space in between the equal sign and the identifier and the actual value. The row will be ignored if the identifier is not known or if there is no equal sign. So additional rows with comments can be introduced. Below such file is shown (grey box) with an explanation of the right.

FILE	Explanation
FISH-quant v0r	
ANALYSIS SETTINGS 01-Jun-2012	
# EXPERIMENTAL PARAMETERS	
lambda_EM=568	Emission wave-length
lambda_Ex=568	Excitation wave-length
NA=1.25	Numerical aperture
RI=1.458	Refractive index
Microscope=widefield	Type of microscope
Pixel_XY=160	Pixel-size [XY] in nm
Pixel_Z=600	Pixel-size [Z] in nm
# GENERAL PROPERTIES	
# FILTERING	**Setting for filtering**
Kernel_bgd=5	Factor for BGD-Kernel
Kernel_psf=1	Factor for PSF-Kernel
# PRE-DETECTION	**Pre-detection**
Detect_Region_XY=2	Size of region around each detected spot in pixel in XY an Z
Detect_Region_Z=2	
Detect_Thresh_int=19	Threshold for intensity
Detect_Thresh_score=2	Threshold for score
Detect_FLAG_reg_smaller=0	Allow different region for pre-detection
Detect_Score=Standard deviation	Type of quality score
Detect_Mode=nonMaxSupr	Type of pre-detection
Detect_FLAG_reg_pos_sep=0	Allow different region for pre-detection
Detect_Region_XY_sep=2	Size of region to perform pre-detection in XY and Z
Detect_Region_Z_sep=4	
Detect_Mode=nonMaxSupr	
# RESTRICTION OF FITTING PARAMETERS	**Restriction of fitting parameters**
sigma_xy_min=0	
sigma_xy_max=511	
sigma_z_min=0	
sigma_z_max=2067	
# AVERAGING	**Size of region used for averaging**
AVG_Region_XY=9	Size of region for averaging in XY and Z.
AVG_Region_Z=9	
AVG_OS_XY=1	Factor for oversampling, i.e. division in how many sub-pixels in XY and Z.
AVG_OS_Z=1	
# THRESHOLDING OF DETECTED SPOTS	**Parameters for thresholding of fitted spots**
SPOTS_TH_sigmaXY_min=114	See below for more details.
SPOTS_TH_sigmaXY_max=187	

Thresholding can be done for each parameter. In order to activate thresholding the allowed minimum and maximum values for this parameter have to be specified. The always have the format **SPOTS_TH_PAR_min** and **SPOTS_TH_PAR_max** where PAR stands for the respective parameter. Supported are **sigmaXY** (Width of the Gaussian in xy), **sigmaZ** (Width of the Gaussian in Z), **amp** (Amplitude of the Gaussian), **bgd** (Local background of the Gaussian), **score** (Quality score of the pre-detected spot), **iter** (Number of iterations in the fit), **resNorm** (Residuals of the fit). Paramters which are not listed will not be thresholded.

Appendix 4 – summary of batch detection: mature mRNA

This file summarizes the results of batch detection for mature mRNA. For each analyzed cells the total number of detected transcripts in the cell are listed before and after thresholding.

Below is such file (grey box). On the left side is an explanation of what the key-words (bold) mean.

Row	**FILE**

```
1  FISH-QUANT    v2c
2  RESULTS OF SPOT DETECTION PERFORMED IN BATCH MODE ON 28-Oct-2013
3
4  COMMENT    Batch detection
5  ANALYSIS-SETTINGS      _FQ_batch_settings_MATURE_131028.txt
6  FILE    CELL    AREA_cell    AREA_nuc    N_total    N_thres_Total    N_thres_Nuc
7  BGH_CY3_01_spots_130925.txt    Cell_1  10341   4859    1     1     1
8  BGH_CY3_01_spots_130925.txt    Cell_2  11090   3749    38    33    16
9  BGH_CY3_01_spots_130925.txt    Cell_3  9204    3240    12    11    5
```

- **Line 1-2** Header row, specifying version number, type of file and date
- **Line 3** is reserved for a user-comment and starts with the key word **COMMENT**
- **Line 4** contains the file-name of the corresponding analysis settings (if it was saved)
- **Line 5** contains a description of the parameters which are listed afterwards.

File	File-name of image
CELL	Identifier of the cells in the image
AREA_cell	Area of the cell as specified by the outline
AREA_nuc	Area of the nucleus (if specified), otherwise contains 0.
N_total	Total number of spots in cell before thresholding
N_thresh_Total	Total number of spots in cell after thresholding
N_thresh_Nuc	Total number of spots in nucleus after thresholding. If nucleus is not defined, the total count in the cell will be repeated.

- **Line 6-end** contains the results of the analysis.

Appendix 5 – file with settings for nascent mRNA quantification

This file defines all parameters describing transcription site quantification. It can be read-in by the batch processing tool. Settings are stored in a simple format. Each property starts with the name followed by a '='and the actual value. There is NO space in between the equal sign and the identifier and the actual value. The row will be ignored if the identifier is not known or if there is no equal sign. So additional rows with comments can be introduced. Below such file is shown (grey box) with an explanation of the right.

FILE	Explanation
FISH-quant v0r	
SETTINGS FOR TRANSCRIPTION SITE QUANTIFICATION	
23-Apr-2012	
# EXPERIMENTAL PARAMETERS	
lambda_EM=568	Emission wave-length
lambda_Ex=568	Excitation wave-length
NA=1.25	Numerical aperture
RI=1.458	Refractive index
Microscope=widefield	Type of microscope
Pixel_XY=160	Pixel-size [XY] in nm
Pixel_Z=600	Pixel-size [Z] in nm
# DESCRIPTION OF PSF AND BGD	
PSF_path_name=	Path for averaged image of mRNA
/Users/guest/Desktop/FISH/results_mature	
PSF_file_name=mRNA_AVG_ns.tif	File-name: averaged image of mRNA
BGD_path_name=	Path-name of BGD image (if specified)
BGD_file_name=	File-name of BGD image (if specified)
PSF_BGD_value=243	Scalar background value
AMP_path_name=	Path name to file with summary of
/Users/guest/Desktop/FISH/results_mature	estimated fitting parameters
AMP_file_name=_FQ_MATURE_SUMMARY_ALL_spots.txt	Name of summary file

fact_os_xy=1
fact_os_z=1
# SETTINGS FOR QUANTIFICATION 	
FLAG_placement=2
FLAG_quality=1
FLAG_posWeight=1
FLAG_crop=1
FLAG_psf=2
FLAG_shift=1
FLAG_parallel=1
N_reconstruct=100
N_run_prelim=50
crop_image_xy_nm=500
crop_image_z_nm=1000
factor_Q_ok=1.5
 	
# SETTINGS FOR BACKGROUND of TxSite 	
FLAG_bgd_local=2
bgd_auto_N_bins=25
bgd_auto_fact_min=4
bgd_auto_fact_max=4
 	
# SETTINGS FOR AUTO-Detection 	
FLAG_auto_detect=0
	Over-sampling factor of averaged image in XY and Z.
	
	See section of options for transcription site quantification for details.

Appendix 6 – summary of batch detection: nascent mRNA

This file summarizes the batch detection for nascent mRNA. For each analyzed transcription site the estimated number of transcript and the standard deviation is reported. As comparison the number of transcripts as obtained with the traditional method of quantification is given as well. Below is such a file (grey box).

Row	**FILE**
1	FISH-quant v0r
2	RESULTS TxSite quantification performed ON 01-Jun-2012
3	COMMENT
4	ANALYSIS-SETTINGS _FQ_settings_NASCENT.txt
5	PROPERTIES OF mRNA
6	mRNA: sigma-XY: 157.421
7	mRNA: sigma-Z: 443.29
8	mRNA: brightest pixel: 168
9	mRNA: sum of all pixels (intensity): 1972
10	mRNA: sum of all pixels (number): 27
11	mRNA: background: 279
12	mRNA: estimated amplitude: 232
13	FILE CELL TS N_IntInt N_PSFsup N_PSFsup_std N_SumPix N_Amp N_MaxInt sigma_xy sigma_z AMP BGD Size_mean[nm] Size_std[nm] BGD_cell TS_PixSum PSF_BgdSum TS_MaxInt
14	BGH_CY3_01__outline.txt Cell_2 TS_1 34 26 2 26 17 18 201.097 603.341 3906.49 503.042 283 15 234 58640 7533 3174

- **Line 1-2** Header row, specifying version number, type of file and date
- **Line 3** is reserved for a user-comment and starts with the key word **COMMENT**
- **Line 4** contains the file-name of the corresponding analysis settings (if it was saved)

- **Line5-12** Summarizes the properties of the individual mRNA molecules used for the quantification: σ_{xy} and σ_z , intensity of brightest pixel, sum of all pixel intensities around brightest pixel, number of pixels to calculate this sum, background of mRNA, and estimated amplitude.
- **Line 13** contains a description of the parameters which are listed afterwards.

File	File-name of image
CELL	Identifier of the cells in the image
TS	Identifier of the transcription sites in the cells
N_IntInt	Number of nascent transcripts: ratio of integrated intensity
N_PSFsup	PSF superimposition approach: mean number of nascent mRNA
N_PSFsup_std	PSF superimposition approach: stdev of nascent mRNA
N_SumPix	Number of nascent transcripts: sum of intensities in pixels around brightest pixel
N_Amp	Quantification with ratio of amplitude of Gaussian fit
N_ratio_int	Quantification with ratio of intensities
Sigma_xy	Fit with 3D Gaussian: σ_{xy}
Sigma_z	Fit with 3D Gaussian: σ_z
AMP	Fit with 3D Gaussian: amplitude
BGD	Fit with 3D Gaussian: background
Size_mean [nm]	Average distance of individual mRNA from center
Size_std [nm]	Stdev of distance of individual mRNA from center
BGD_cell	Background of cell used in the PSF superimposition approach
TS_PixSum	Sum of all pixel intensities around brightest pixel. Used for the quantification N_SumPix.
PSF_BgdSum	Sum of all pixels for background in averaged image of mRNA. Same range as for TS_PixSum.
TS_MAX_INT	Maximum intensity of transcription site (before BGD subtraction)

- **Line 14-end** contains the results of the analysis.

Appendix 7 – summary of batch detection: nascent and mature mRNA

This file summarizes the batch detection for mature and nascent mRNA. For each analyzed cell the number of mature mRNA and the number of nascent mRNA at each of the associated transcriptions sites is reported. Below such file is shown (grey box).

Row	**FILE**
1	FISH-quant v0r
2	RESULTS TxSite quantification performed ON 01-Jun-2012
3	COMMENT Batch detection
4	ANALYSIS-SETTINGS _FQ_batch_settings_MATURE_120601.txt
5	FILE CELL N_MATURE_Total N_MATURE_Nucleus
6	N_NASCENT_FOR_EACH_TS_IN_CELL
7	BGH_CY3_01__outline.txt Cell_1 1 0
8	BGH_CY3_01__outline.txt Cell_2 31 12 4
9	BGH_CY3_01__outline.txt Cell_3 13 5

- **Line 1-2** Header row, specifying version number, type of file and date
- **Line 3** is reserved for a user-comment and starts with the key word **COMMENT**
- **Line 4** contains the file-name of the corresponding analysis settings (if it was saved)

- **Line 5** contains a header for the subsequent rows

File	File-name of image
CELL	Identifier of the cells in the image
N_mean_MATURE_Total	Number of mature mRNA in cell (always reported)
N_mean_MATURE_Nucleus	Number of mature mRNA in nucleus (always reported). If nucleus is not define, total number in cell is reported.
N_NASCENT	Number of nascent mRNA at each transcription site.

- **Line 6-end** contains the results of the analysis.

In the example above three cells were in the image with a total of three cells were analyzed. One cell contained one transcription site, the other two none.

Appendix 8 – specifying different detection thresholds in batch mode

This file can be used when different detection thresholds are applied to different files in the batch detection. Each file is specified with the file-name that should be the same name as used in the file list specified to be processed (without the extension), followed by the threshold for the minimum intensity, and the minimum quality score (ORDER IS IMPORTANT, values are tab-delimited). The name has to match the name that you specify in the batch mode (either the outline definition file or the actual image file). FISH-quant will use the default settings if no exact match for the name can be found!

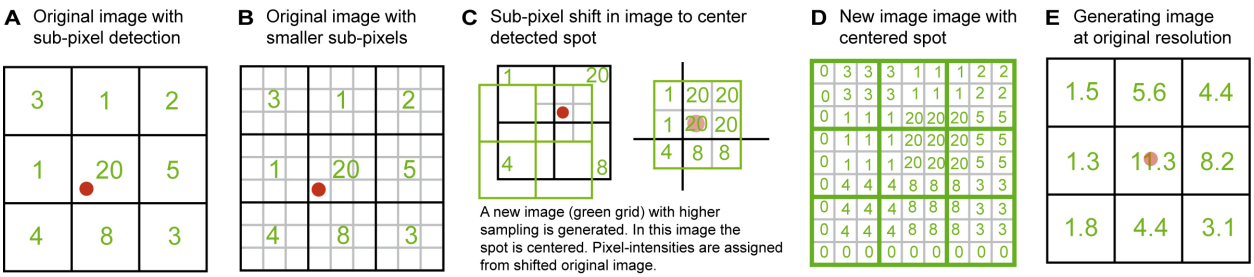
ROW	FILE
1	Name TH_predect TH_score
2	Img1.tif 490 50
3	Img2.tif 500 45

- **Line 1** Header row, specifying rows
- **Line 2-3** Detection thresholds for the different files

Appendix 9 – averaging images at subpixel resolution

The cartoon illustrates how images can be averaged under consideration of the subpixel localization.

Assuming that the imaged molecules are diffraction limited in size we will obtain a slightly different image depending on where the fluorophore(s) are located with respect to the pixelated grid of the image. The sub-pixel detection accuracy of our fitting approach can then be used to obtain an image at higher resolution with improved quality. The basic idea behind this approach is outlined in the image below



We know the center of each spot with sub-pixel accuracy (Red spot in panel A). We can then divide the pixels in the original image into smaller sub-pixels where each sub-pixel has the same intensity as the original pixel (Panel B). Then we generate a new image (Panel C, green grid) where we shift the original

image such that the identified center is truly center in the new grid (Panel D). We can then use this shifted image to calculate an image in the original resolution by averaging all values of the sub-pixels (Panel E).

12. References

1. Piotr's Matlab Toolbox. at <http://vision.ucsd.edu/~pdollar/toolbox/doc/>
2. Bio-Formats | LOCI. at <http://www.loci.wisc.edu/software/bio-formats>
3. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9,** 676–682 (2012).
4. Sage, D., Prodanov, D. & Ortiz, C. MIJ - A Java package for bi-directional communication and data exchange from Matlab to ImageJ. at <http://bigwww.epfl.ch/sage/soft/mij/>
5. Birngruber, E., Donner, R. & Langs, G. matVTK - 3D visualization for Matlab. *Proceedings of the MICCAI 2009 Workshop on systems and architectures for CAI*
6. Zhang, B., Zerubia, J. & Olivo-Marin, J.-C. Gaussian approximations of fluorescence microscope point-spread function models. *Appl. Opt.* **46,** 1819–1829 (2007).
7. MATLAB — Bio-Formats 5.0.6 documentation. at <http://www.openmicroscopy.org/site/support/bio-formats5/users/matlab/index.html>
8. Belmont, A. S. & Straight, A. F. In vivo visualization of chromosomes using lac operator-repressor binding. *Trends Cell Biol.* **8,** 121–124 (1998).
9. Henriques, R. *et al.* QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nat. Methods* **7,** 339–340 (2010).
10. Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Meth* **5,** 877–879 (2008).
11. Thomann, D., Rines, D. R., Sorger, P. K. & Danuser, G. Automatic fluorescent tag detection in 3D with super-resolution: application to the analysis of chromosome movement. *J Microsc* **208,** 49–64 (2002).