Image analysis and quantification guide for

Absolute quantification and single-cell dose-response of cytosolic siRNA delivery

Hampus Hedlund, Hampus Du Rietz, Johanna M Johansson, Wahed Zedan, Linfeng Huang,

Jonas Wallin & Anders Wittrup

March 2022

About this guide

This is a short guide for the image analysis steps used for the quantification of cytosolic delivery and dose-response of siRNA presented in the paper. We provide an example dataset, CellProfiler pipeline and custom MATLAB code with a graphical user interface (cytosolQuant) to try the various steps of the approach and serve as a starting point for adaptation to other datasets. For additional background and analysis considerations, we refer to the main paper and the Methods section.

The example dataset and MATLAB code is available via figshare.com. The MATLAB code can also be accessed via github.com/WittrupLab/cytosolQuant. The image data is provided as a zip-file. A folder (exampleEXP01) containing the completed analysis of the dataset is also provided, to serve as a reference during testing. This folder also includes the data generated by CellProfiler.

The example dataset includes time-lapse images from an experiment were HeLa-d1eGFP cells were treated with lipoplexed siGFP-2 at a ratio of 2:4 pmol:µL siRNA to Lipofectamine 2000. Each image-file name contains metadata such as frame number, *z*-plane and channel and can be interpreted as:

Metadata	Location
Frame number	rawimages_t**XXX**zXcX_ORG
Z-plane	rawimages_tXXXz**X**cX_ORG
Channel	rawimages_tXXXzXc**X**_ORG

The acquired channels are in turn denoted:

Assigned name	Image sets
d1eGFP	rawimages_tXXXzXc**2**_ORG
nuclei	rawimages_tXXXzXc**3**_ORG
siRNA	rawimages_tXXXzXc**4**_ORG

Acknowledgement

This work was carried out in the labs of Anders Wittrup and Jonas Wallin, Lund University, and supported by funding from the Swedish Society for Medical Research (SSMF).

1. Download and cytosolQuant setup

Make sure that MATLAB, CellProfiler (v. 2.2.0) and Fiji are installed on the computer.

☞ Step-by-step

1. Download rawImages.zip (example dataset), Pipeline.cpproj (CellProfiler pipeline), and cytosolQuant.zip (analysis folder structure with completed example experiment folder and MATLAB code) at figshare.com.
2. Unarchive cytosolQuant.zip and rawImages.zip.
3. Run the cytosolQuant application by clicking cytosolQuant.mlapp in .../cytosolQuant/Processing tools/Matlab. The App opens in a separate window from the main MATLAB software.
4. In the Settings tab, provide the full path to the cytosolQuant folder (Home folder).
5. In the MATLAB main window, set the current directory to Volume/User/.../cytosolQuant/Processing tools/Matlab.

The MATLAB code uses a well-defined folder architecture. In the provided files, the Home folder is named cytosolQuant. This directory contains the subfolders

- Data repository containing all data in a hierarchical fashion, where acquisitions are organized in the folder of their corresponding experiment. The acquisition folder contains all the data associated with the specific dataset. The folder exampleEXP01 contains all data and files (except the provided example raw images) generated during the completed analysis up until modeling of siRNA release and knockdown.
- Processing tools containing tools necessary for processing and analysis, including the subfolder Matlab with all MATLAB functions and code.

2. Setting up experiment parameters (cytosolQuant)

Experiment parameters should be specified in the Experiment section in the App. To activate an experiment tab for processing, the top left box Activate experiment tab must be ticked.

A maximum of ten experiments can be processed during the same run. The following information should be entered:

- Experiment ID is the name of the experiment with appropriate syntax.
- N Time Series specifies the number of acquisitions in the experiment. The value only needs to be specified to execute the Create Folders operation.
- Channel A, B and C specifies the name of the channels in the dataset. The order of the channels entered here need to match the order of the channels in the dataset (c1, c2, c3). Enter any name or choose from the list.

- x Dim, y Dim, z Dim is the spatial dimension of the datasets (pixels).
- Control positions is used to specify which of the acquisitions (positions) in the experiment that were used as internal controls (no treatments).
- Include Time Series specifies if all acquisitions in the experiments should be processed, or only acquisitions that are marked (selected) by temporarily adding a hashtag last in the filename of the folder in the Data repository. Changes to this setting (and related options in the Processing tabs) is not required for the example dataset.

☞ Step-by-step

1. In the first tab in the Experiment section, check the Activate experiment tab box to activate the tab. As experiment name, enter exampleEXP02.
2. Use the prespecified parameters for the number of subfolders (acquisitions) for the experiment (N Time Series), channel names and image properties.

3. Adding experiment folders and data (cytosolQuant)

The App is used to generate the appropriate folder structure of single or multiple experiments. After providing experiment information in the Experiments section, press the Create Folders button in the Pre-processing tab to create all folders with the correct structure. Experiment names must start with a letter, and can only contain letters, numbers, underscores or dashes.

☞ Step-by-step

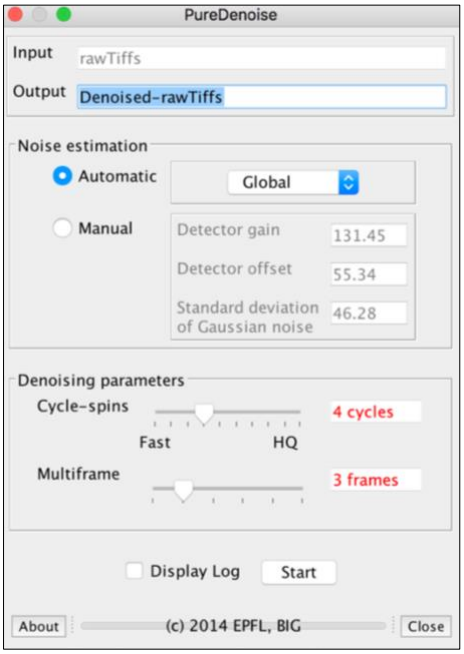
1. Click the Create Folders button.
2. Manually transfer the downloaded raw image tiff-files to cytosolQuant/Data repository/exampleEXP02/exampleEXP02_TS01/rawTiffs, with no additional subfolders.
3. No control dataset is provided with this short guide. Without any selection in the Control positions box, click the Save control postions button.

4. Image denoising (Fiji)

Denoising is performed to increase image signal-to-noise, to improve the reliability of segmentation, object identification and tracking in later steps. We have used the ImageJ plugin PureDenoise. (Luisier F, Blu T, Unser M. *IEEE Trans Image Process.* 2010a. Image Denoising in Mixed Poisson-Gaussian Noise).

☞ Step-by-step

1. Run Fiji.

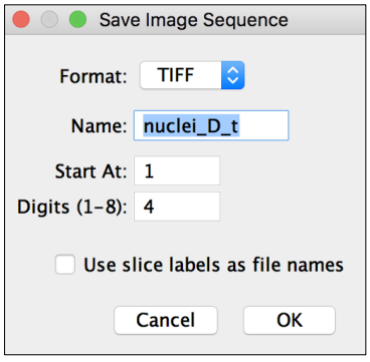


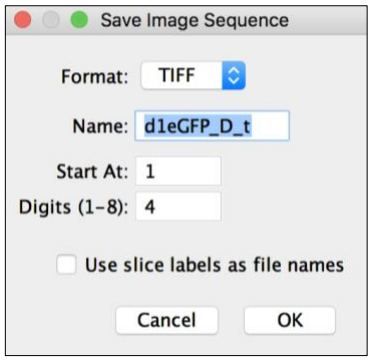
2. Import the image sequence to be denoised. Choose either rawimages_tXXXz1c**2**_ORG (d1eGFP channel) or rawimages_tXXXz1c**3**_ORG (Hoechst33342 channel) located in the folder rawImages. awimages_tXXXz1c**1**_ORG (siRNA channel) and rawimages_tXXXz1c**4**_ORG (siRNA channel) should not be denoised.

3. Open PureDenoise under Plugins.

4. Run denoising using the default setting (shown to the right).

5. Save denoised image sequence to exampleEXP02/exampleEXP02_TS01/denoisedGFP or exampleEXP02/exampleEXP02_TS01/denoisedNuc depending on the channel that was denoised. Name them d1eGFP_D_t or nuclei_D_t and use settings as shown below.





6. Repeat step 2–5 until both channels have been denoised.

5. Cell tracking and fluorescence quantification (CellProfiler)

CellProfiler is used for cell segmentation, tracking and fluorescence quantification. The custom analysis pipeline is highly adapted to suit the analysis needs and the quality and dimensions of the data.

☞ Step-by-step

1. Start CellProfiler and load pipeline Pipeline.cpproj.
2. Open the Images module and import the denoised d1eGFP and nuclei images, and the raw tiff images from exampleEXP02/exampleEXP02_TS01, by dragging the folders denoisedGFP, denoisedNuc and rawTiffs into the file list window.
3. Open the Metadata module and press Update in the window below module settings to verify that metadata has been correctly extracted from the imported images. Metadata on frames should be extracted from all images. Metadata on channel and stack is only extracted from the image files named rawimages_tXXXzXcX_ORG.tif.
4. Open the NamesAndTypes module and press Update in the window below module settings to verify that names have been assigned to the correct image sets as indicated in the table below.

Assigned name	Image sets
d1eGFP	rawimages_tXXXz**1**c**2**_ORG
nuclei	rawimages_tXXXz**1**c**3**_ORG
siRNA_z1	rawimages_tXXXz**1**c**4**_ORG
siRNA_z2	rawimages_tXXXz**2**c**4**_ORG
denoisedGFP	d1eGFP_D_tXXXX
denoisedNuclei	nuclei_D_tXXXX

5. Open the ExportToSpreadsheet module and make sure all measurements are selected for export by pressing the button Press button to select measurements and select All at the top of the list.
6. Start the analysis by pressing the Analyze Images button in the bottom left corner.

7. After analysis, move the exported Excel spreadsheets to the following folders in file directory exampleEXP02/ExampleEXP02_TS01/cellprofilerOutput/datasheet as indicated in the table below. Note, the spreadsheet "Image" must be exported as .xlsx or similar (in Excel) for downstream analysis to work as intended.

Spreadsheet	Folder
Image	backgroundIntensity
Cells_z1z2	cellFullsiRNA
MaskedCells	CellMaskedsiRNA
identifiedNuclei	identifyNuc

To skip the CellProfiler analysis and directly proceed with the subsequent steps in the guide, simply copy the contents in exampleEXP01/exampleEXP01_TS01/cellprofilerOutput/datasheet to the corresponding folder in exampleEXP02.

6. Primary processing (cytosolQuant)

After setting up the experiment parameters in the App, adding the raw image tiff-files to the rawTiffs folder and the CellProfiler output as outlined above, the next step is Primary processing. This can be executed with multiple experiments activated in the Experiments section, or only one active experiment at a time. Several operations are carried out. Images are sorted into new folders and resampled. Data from individual cells are imported from the CellProfiler datasheets and saved separately. Release events, mitosis, cells lacking d1eGFP expression are detected in siRNA or GFP traces.

☞ Step-by-step

1. Press Primary processing under Processing tab.

7. Creating ROIs of events (cytosolQuant)

After object or event identification, the relevant data for further analysis of individual objects can be extracted as event ROIs of the full dimensional dataset, with some associated whole-cell data and metadata. In the App, options for creating the ROIs should be specified accordingly:

- Extend ROI specifies the lateral *xy*-dimension of the ROI, indicated as the pixel distance from the center given by the event location (default 125 pixels).
- Times Pre specifies the number of time points before the indicated event to include in the ROI (default 10).
- Times Post specifies the number of time points after the indicated event to include in the ROI (default 20).

☞ Step-by-step

1. Select the Event ROIs tab.
2. Use all prespecified parameters.
3. Press Create Event ROIs.

8. Create event panels (cytosolQuant)

To make inspection and validation of events easier and less time consuming, ROIs of cells with events can be put together in a new collage – here referred to as event panels. A path to a folder where the event panel is saved should be provided in the Export Folder field. Typically, this folder can be the first acquisition folder of the experiment processed. If multiple experiments are included in the event panel, a single panel will be generated and exported to the location specified. Event panels can be selected to include events that are not validated (default), validated, false positive or all events regardless of validation status. Also, panels can be specified to include events that appearing first (primary events), second (secondary events) etc. in the cell.

☞ Step-by-step

1. Select the Event Panels tab.
2. Use all prespecified parameters.
3. Provide the full directory path of the export folder (Volume/User/…/exampleEXP02/ExampleEXP02_TS01).
4. If desirable, change the name of the panel. Panels with the same name will not be overwritten.
5. Create ROI panel by pressing the button Create ROI Panel in the bottom right corner.

9. Event validation (Fiji, cytosolQuant)

An algorithm was used to achieve automatic detection of siRNA release events. The tuning of the detection was performed so that the number of false negative events (*i.e.* undetected true release event) was minimized. Due to the nature of the data and the biological conditions, this in turn made the algorithm more prone to flag non-release related intensity fluctuations as release events (referred to as false events). To correct such errors, we implemented a manual release event validation.

☞ Step-by-step

1. Run Fiji and import the Event panel using Import > Bio-Formats. A panel with multiple detected events will load. Channels are presented in the following order: d1eGFP (c1), nuclei (c2), siRNA-AF647 (c3).

1. Choose Image > Color > Make Composite and adjust the LUT and brightness to visualize the cells. Set brightness of siRNA channel (c3) to min = 0 and max = 1000. This range can usually give a first impression of the detected release events.

2. Open the ROI manager (Analyze > Tools > ROI Manager).

3. Visually evaluate the event detected in each cell. False positive events, events detected after release has already occurred (if primary events) and cells or events with clearly aberrant appearance should be detected.

4. The event status is marked using the Point Tool. Adding one point located within the event (image) ROI to the ROI Manager will mark the event as false positive. Adding two points located within the event ROI to the ROI Manager will disqualify the event for further analysis. Release events identified correctly are not marked, and will automatically be treated as true positive events. It is important that the ROIs are marked in the d1eGFP channel (first channel in Fiji) for the downstream analysis to work properly.*

5. After evaluating all events, save the completed ROI list by cliking More > List > Save as…. Save list as a csv-file in the Event panel folder. Name the file falseEvents.csv.

6. In the cytosolQuant App, open the Event validation tab.

7. Enter the Event panel folder path (where the falseEvents.csv file was saved) and press Add new validation data.

8. Create a new panel as described in the previous section. The new panel will contain the next event detected in cells that was previously marked as having had a false event. True events that were already validated and disqualified cells will not be included. Repeat Step 1–7 above. If the Event panel only contains valid events, update the Event panel folder path in the Validation data tab in cytosolQuant, select No false events present in panel and press Add new validation data.

9. If all primary events are validated, continue with secondary events until they have been validated as well. A new panel with secondary events can be created by selecting Include events appearing 2nd in the cell. This validation step is not essential for completing this guide, but is necessary to obtain reliable knockdown kinetics from primary events and mask data after the second event, which is vital when using the data for modeling.

* When used in MATLAB, the variables xVar, yVar, EvalxVar and EvalyVar in the falseEvents table must be indexed from the right position. Different Fiji versions/options will change the index of the column containing the needed variable, which in that case must be updated for the correct result.

10. Secondary processing (cytosolQuant)

When siRNA release events have been identified and validated, in the next step reference measurements consisting of image stacks of 0 nM and 1000 nM siRNA suspension are used for normalization of the siRNA data. Also, a correction of the illumination across the entire field-of-view is applied.

☞ Step-by-step

1. Reference measurement images are located in cytosolQuant/Data repository/exampleEXP01/exampleEXP01_TS01/referenceMeasure. Copy the entire folder and paste in the corresponding folder exampleEXP02_TS01.
2. In the Processing tab, click the Secondary processing button.

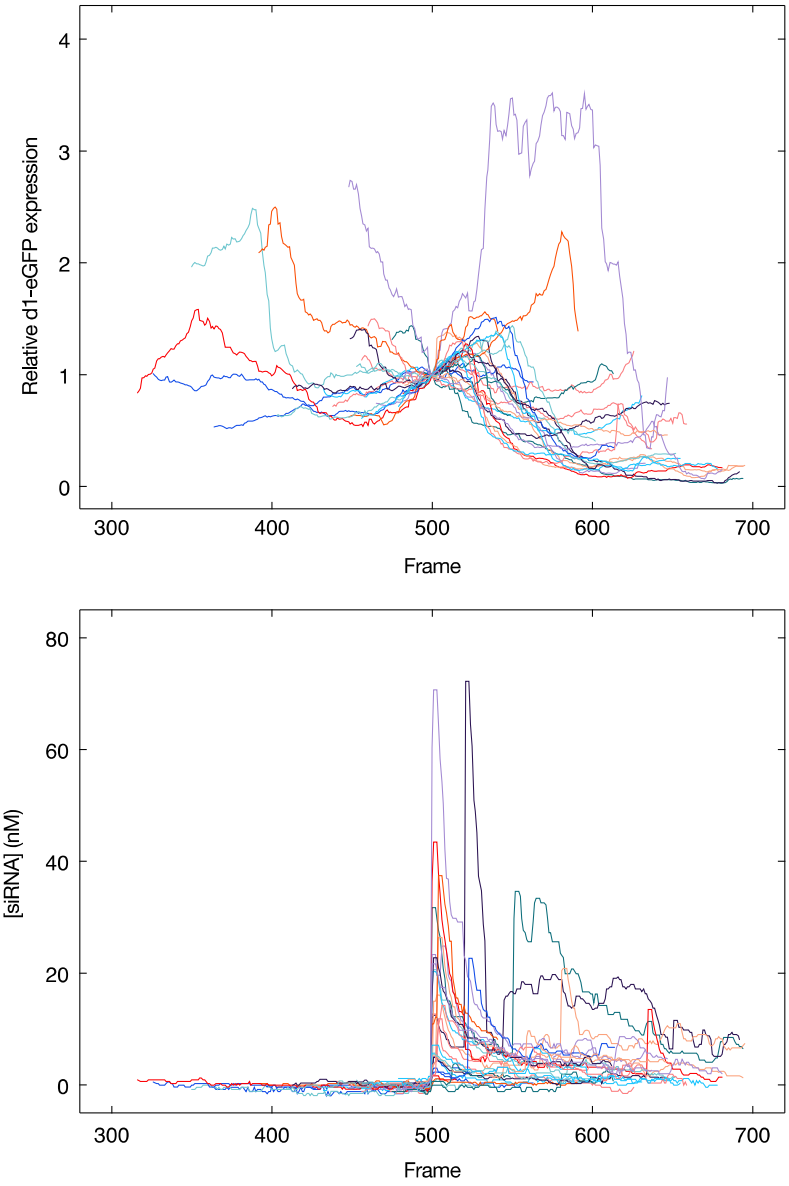
11. Collect data (cytosolQuant)

The final step of the processing is collecting the relevant output data into a single datasheet. The data is also corrected for bleaching using traces of cells where no events were detected (or control cells when available).

Data is collected in the structure array collectedData in the MATLAB Workspace. Data is collected both as time shifted and with standard time. In the time shifted data, frame 500 represents the time-point where the release event was detected. The data is also presented with traces masked from the point where a second release event is detected (maskedData), which is typically used. Note that in this demo no secondary events were validated and hence data is not masked in this dataset. In the array, data is presented as rawTraces (raw measurements), mfTraces (median filtered values), mfBcTraces (median filtered and siRNA-background corrected), bcTraces (siRNA-background corrected). The channel siRNA is the siRNA channel measured with the nucleus masked, while siRNAF is the measurement in the entire cell mask (only used for event detection). Typically, the time-shifted, median-filtered and background corrected eGFP and siRNA traces are used for post-processing analysis and modeling.

☞ Step-by-step

1. Select the Collect tab.
2. Use the prespecified parameters and Release events, primary.
3. Click Collect Data.
4. The collected data will appear in the MATLAB Workspace as collectedData.



The data output (collectedData.timeShift.maskedData.mfBcTraces.eGFP and collectedData.timeShift.maskedData.mfBcTraces.siRNA) shows the relative d1eGFP fluorescence intensity, normalized per cell to the value at frame 500 (top), and the measured cytosolic siRNA concentration (bottom). Traces are shifted in time so that the release event is detected at frame 500.

11. Modeling cytosolic siRNA release and knockdown (MATLAB)

Refer to the main paper and Supplementary information for a detailed description of the siRNA release and knockdown modeling. For modeling, the d1eGFP data should not be normalized, as in the example above. Separate example data (including additional data to the experiment analyzed in this guide) is available via github.com/WittrupLab/cytosolQuant but is also included in the cytosolQuant folder under siRNA and knockdown modeling together with the MATLAB code needed. D1eGFP data can also be collected without normalization by unchecking Normalize d1eGFP traces in the Collect tab in the App. The graphical user interface is not used for the analysis and modeling outlined below.

☞ Step-by-step

siRNA quantification model

1. Change current directory to .../code/siRNA
2. Open inputEstimationRNA.m. The script will fit three siRNA quantification models to cytosolic siRNA measurements.
3. Press button during run for next cell. Set buttonPressFlag = 0 for uninterrupted run. Total run time is 10 min on standard computer, ~10 s for a single cell.

d1eGFP knockdown and siRNA release magnitude

4. Change current directory to .../code/QuantilePlots
5. Open plotReleaseSortedeGFP.m. The script will plot eGFP responses for quartiles of cells treated with siGFP-1 and with quantification R^2>0.75.

d1eGFP knockdown model and siRNA release magnitude

6. Change current directory to .../code/eGFP
7. Open plotGFPmeans.m. The script will plot eGFP downregulation model estimates relative to individual traces of eGFP expression with groups divided by release magnitude quintiles