Data Analysis Protocol for NIS-Seq Optical Pooled Screens
v1.1 I 8/24
Caroline I. Fandrey, Schmid-Burgk Lab, University of Bonn
http://jsb-lab.bio/opticalscreening/

Data and Data Preparation:

- Single files of phenotype data (one file per channel/tile/well)
    - If z-stacks were acquired, combine z-stacks of same tile to mean projection

- Single files of NIS-Seq cycles
    - Combine images of all NIS-Seq cycles in one folder per well acquired

Analysis steps:

Cell Pose

The CellPose algorithm (Stringer et al., Nat Methods. 2021) is used to translate membrane or nuclear staining data into cell masks. Later, these masks define the exact area of the nucleus and where the outer membrane of one cell ends and the next cell starts. By this, we can precisely assign NIS-Seq sequences to the nucleus of one cell and calculate fluorescent signal correlation between nucleus and cytosol.

To run CellPose, the following command is used:

cellpose --dir ***/user/hard drive/experiment folder/target folder/*** --pretrained_model ***model*** --diameter ***xx*** --use_gpu --verbose --no_npy --save_tif

with the following parameters:

- nuclei model for nuclei masks of NIS-Seq images acquired with 20x objective
    - HeLa cells: 30 μm diameter
    - THP-1 cells: 30 μm diameter
- nuclei model for nuclei masks of phenotype images
    - HeLa cells (20x objective): 30 μm diameter
- cyto2 model for membrane masks of phenotype images
    - HeLa cells (20x objective): 50 μm diameter
    - THP-1 cells (10x objective): 20 μm diameter

1) Analyze NIS-Seq raw imaging data

→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "1. Analyze NIS-Seq raw imaging data"
→ Enter experiment name and well ID
→ If necessary, adjust number of cycles imaged in the experiment

NIS-Seq Analysis v1.0
JSB lab 2020-2024

1. Enter experiment or well name:

2. Enter number of cycles:
14

3. Load In-situ images (4 channels, 2048x2048, 16bit, sorted):
Choose files No file chosen

4. Load nuclear masks (Generate with CellPose, 1 channel, 2048x2048, 16bit, sorted):
Choose files No file chosen

5. Load or calculate NIS-Seq cycle alignment:
Load: Choose file No file chosen (tab delimited, x (px) - y (px), no header) (re-load)
Or calculate: Calculate alignment (7 seconds per image)

6. Load or detect spots:
Load: Choose file No file chosen (tab delimited, tile - x (px) - y (px), no header) (re-load)
Or calculate:
Brightness threshold: 350 Detect spots

7. Load compensation Matrix (no header):
Choose file No file chosen (re-load)
compensation_matrix_NextSeq2000_jsb-lab_2022.txt

8. Perform sequence calling:
Optionally Limit analysis to tiles (either from or range first-last, counting from 0):
Start sequence calling

9. Filter spots to match reference library:
Choose file No file chosen (tab delimited, gene - sequence, no header) (re-load)
Download Brunello human sgRNA library and scrambled control
Filter

10. Determine maximum NIS-Seq intensity per nucleus across cycles and channels:
Start measurement

11. Collapse sequences to nuclei
Nucleus intensity threshold: 700000 au
Minimum relative intensity of top sequence: 66.7 %
Assign library-matching sequences to nuclei

1.1: Align in-situ sequencing data by channel 2 (nuclei)

In this step all acquired NIS-Seq cycles will be precisely aligned by the nuclei of the cells. Thereby, base calling can be performed for each cell by transforming the sequential fluorescent signals in the same NIS-Seq spot into a sequence. The output *txt* file contains the x-y pixel-shift of each tile over all cycles based on cycle 1

→ at step *3. Load In-situ images* click “Choose files” and load all NIS-Seq images of one well (all cycles, all channels)
→ at step *4. Load nuclear masks* click “Choose file” and load masks generated with CellPose from the nuclei of the first NIS-Seq cycle
→ at step *5. Load or calculate NIS-Seq cycle alignmen*t load previous alignment or click “Calculate alignment (7 seconds per image)” and wait until all cycles are aligned

- Note: this is one of the longest analysis steps

1.2: Spot detection

Based on the first 3 cycles, the software detects the exact localization of sequencing spot signals for each cell. The threshold can vary between experiments or even wells and is dependent on the brightness of the spots and the background of the images. If a new analysis is started it is always useful to try a snippet of 10 tiles with some spot thresholds and compare the resulting spot file with the raw images and the spots you see by eye. The output file of this step defines which pixel location will be used to generate a sequence.

Once the alignment step is done:
→ at step *6. Load or detect spots* load previous spot file or start new spot detection by clicking the button “Detect spots”

1.3: Base calling

Next, a compensation matrix which is the same file for every experiment run with Illumina NextSeq chemistry is provided. It defines which nucleobase corresponds to which fluorescent channel and thereby is needed to transform the sequential “light signals” into a DNA sequence. This step creates several sequence files (chopped at 100 MB size) which need to be combined in the end to yield one continuous list of sequences for each tile of the whole well (e.g., there are 468 tiles acquired per 24-well using the 20x objective and 1x tube lens → first column starts at 0 and counts up to 467). The combined output *txt* file contains all sequences that were found in the NIS-Seq data.

→ at step *7. Load compensation Matrix”* load the provided compensation matrix “compensation_matrix_NextSeq2000_jsb-lab_2022.txt”
→ at step *8. Perform sequence calling”* click the button “Start sequence calling”

1.4: Filter Spot Sequences to match reference library

In this step all sequences that do not match the reference barcode library will be removed. Therefore, the generated sequences from step 1.3 and a reference barcode dictionary will be merged by overlapping sequences.

→ in step *9. Filter spots to match reference library* click “Choose file” and load a reference or scrambled control library and click the button “Filter”

1.5: Determine spot intensity over whole nucleus
Using the NIS-Seq nuclear masks, the spot intensity is calculated over the whole nucleus area of each cell. This table will later be used to set a threshold to cells with a defined minimal spot signal over its whole nuclear area. By this, cells with fluorescent background spots or artefacts can be eliminated.

→ at step *10. Determine maximum NIS-Seq intensity per nucleus across cycles and channels* click the button “Start measurement”

1.6: Assign NIS-Seq spots to nuclei
This step assigns the created sequences to the correct cell by the xy coordinates of the nuclei.

→ once measurement has run through in step *11. Collapse sequences to nuclei,* if necessary, adjust the nucleus intensity threshold and click the button “Assign library-matching sequences to nuclei” to combine all analysis steps

The final output file contains a list of cells with a library matching sequence detected and an overall spot intensity above the threshold set in the last step.

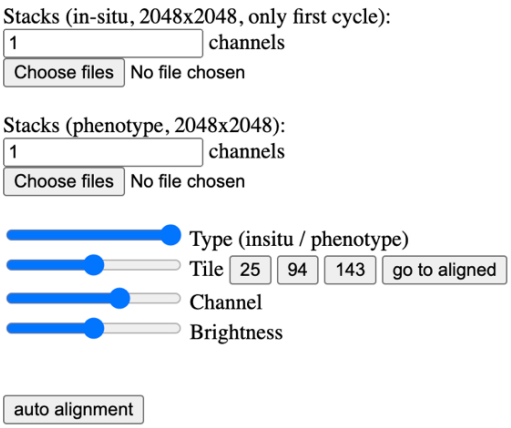
2) Mapping of Phenotype to in-situ images

In this step of the analysis, it is precisely traced back which cell from the phenotype images is corresponding to which cell in the NIS-Seq images – even if the magnification of the microscope objectives is different.
The tool makes use of the nuclear staining and the shape of the nuclei, which allow accurate tracing cells over different imaging modalities like a fingerprint.

For imaging of phenotype and NIS-Seq with the same objective (if the different objectives were used, see anaylsis starting at step 3.1b)

3.1a: one step (same objective)

→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "2. Mapping of Phenotype to Insitu Images: one-step (same objective)"
→ Load NIS-Seq data of only the nuclei in the first cycle
→ Load the phenotype images of the nuclei
→ Press “auto alignment”

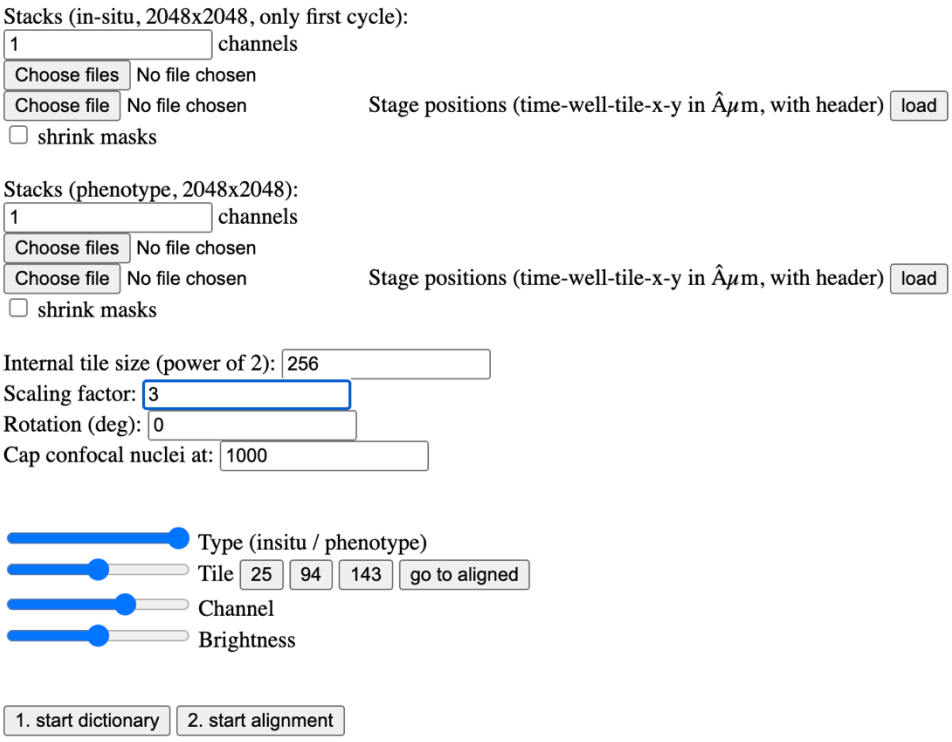


3.1b: step 1 (course)

→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "2. Mapping of Phenotype to Insitu Images: step 1 (course)"
→ Load NIS-Seq data of only the nuclei in the first cycle
→ Load the tile positions file (automatically generated by the microscope software when imaging is started) for the NIS-Seq imaging settings
→ Load the phenotype images of the nuclei
→ Load the tile position files for the phenotype imaging microscope settings

→ Set the scaling factor (e.g. “3” if phenotype was imaged with 40x 1.5x lens (=60x) and in-situ with 20x 1x lens)
→ Press “1. start dictionary “(takes a few minutes)
→ Press “2. Start alignment”

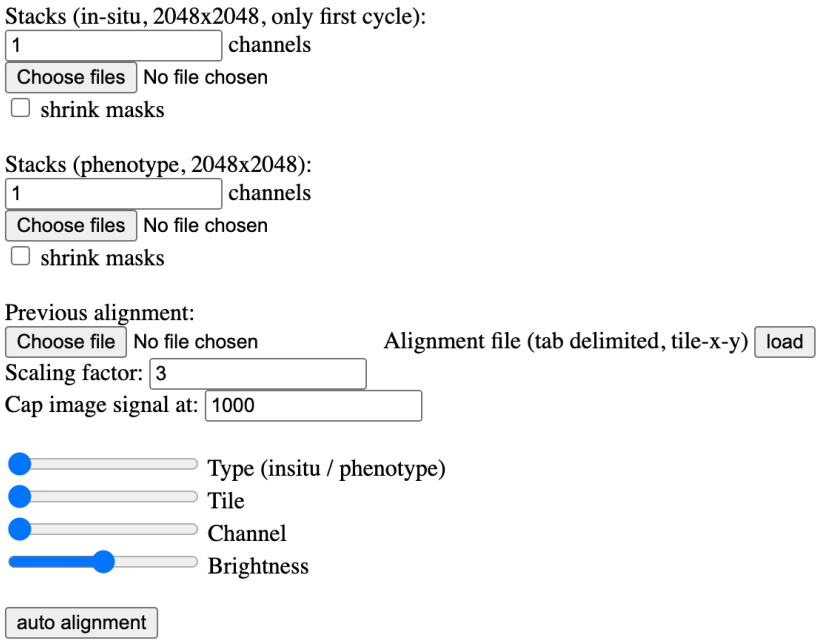
The output file is the first, rough alignment of phenotype to in-situ images but needs to be improved in the next step



3.2: step 2 (refine)

→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "2. Mapping of Phenotype to Insitu Images: step 2 (refine)"
→ Load NIS-Seq data of only the nuclei in the first cycle
→ Load the phenotype images of the nuclei
→ Load the previous alignment file from step 3.1b
→ Set the scaling factor (e.g., “3” if phenotype with 40x 1.5x lens (=60x) and in-situ with 20x 1x lens)
→ Press “auto alignment”

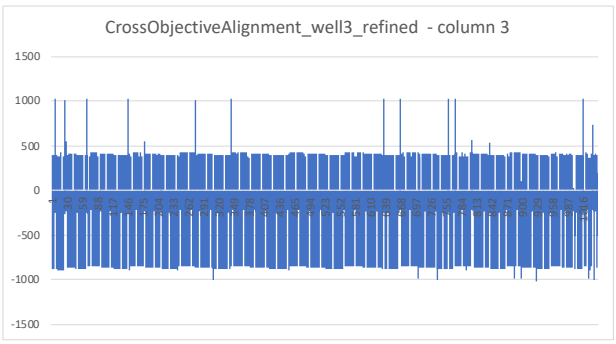
The output file is the refined alignment of phenotype to in-situ images which contains information about the pixel shift between two tiles of the same picture frame. The file needs to be cleaned from outliers in the next step.



3.3: Clean Alignment (for imaging with different or same objective)

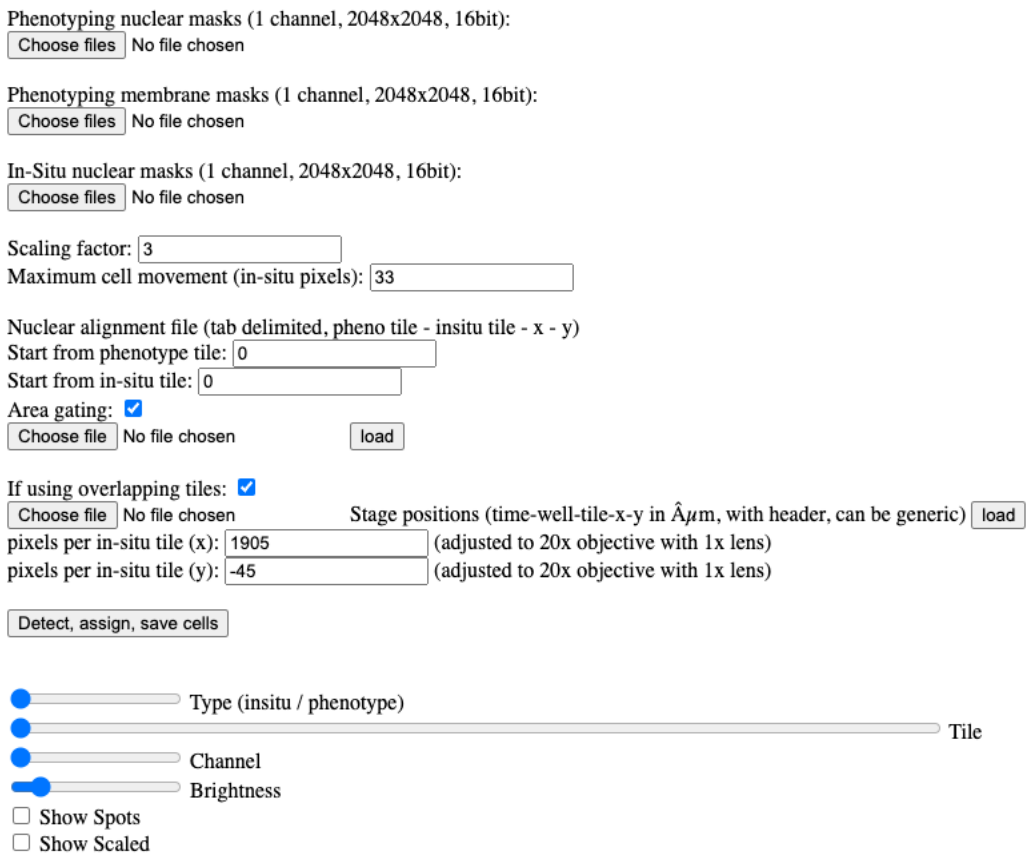
→ Open the file in Excel. Select the 3rd column to plot as bar chart
→ most of the peaks are of a similar height but there are some outliers
→ determine outliers and adjust in excel

- option A: remove tiles that were not properly aligned between NIS and phenotype by just deleting the rows with outlier values
- option B: determine outlier values and interpolate with values of surrounding tiles



3) Link Nuclei between Phenotype and In-Situ Images

In the last step the assigned sequences from the NIS-Seq Data is mapped to the phenotype data.



→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "3. Link Nuclei between Phenotype and In-Situ Images"
→ Load the nuclear masks and membrane masks from the phenotype data

- Note: In the THP-1 screens, the membrane staining was used as a proxy for the cell nuclei, as the area between the membrane and the cell nucleus does not differ significantly.

→ Set the scaling factor (e.g., "3" if phenotype with 40x 1.5x lens (=60x) and in-situ with 20x 1x lens)
→ For Nuclear alignment file, load the refined and filtered file generated for the nuclear mapping between phenotype and in-situ in step 3.3 and press "load"

- Area gating can optionally be enabled if the cells are dense. Here, the nuclear area is used to support correct matching.

→ If NIS-Seq and phenotype images were acquired with different objectives, load the stage position file generated during NIS-Seq imaging and press "load" to enable choosing matching cells also from neighboring tiles.

- If all images were acquired with the same objective, this step can be disabled

→ Start the tool by clicking "detect, assign, save cells"

- The output file is an assignment table of all cells that where found in the NIS-Seq data analysis and the phenotype acquisition

4) Quantify Phenotypes

Depending on the phenotype of interest, different analysis steps were performed.

5.1: Quantify Correlation Phenotypes

To determine the translocation of the fluorescence signal from the cytosol to the nucleus, the membrane masks and the raw images (membrane staining, nuclear staining and fluorescent, translocating protein) must be loaded into the software.

→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "4. Quantify Correlation Phenotype”
→ Load cell masks and raw phenotype images
→ Click "start anaylsis"

A correlation value for the fluorescent protein between cytosol and nucleus is calculated for each cell and entered in the output file.

5.2: Quantify Specking Phenotypes

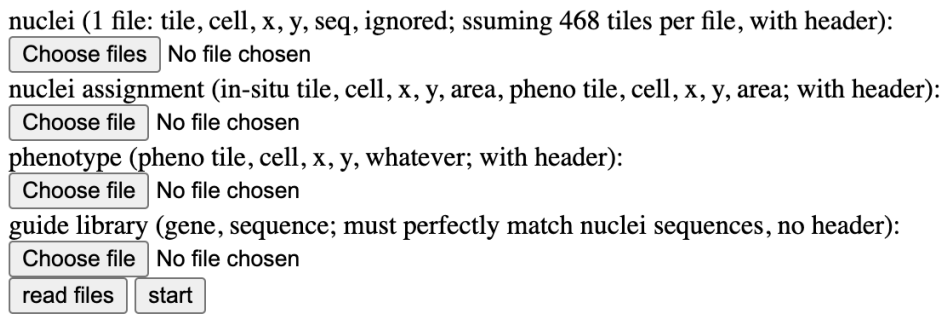
To determine whether bright spots of fluorescent signal (e.g. ASC specks) are formed upon activation, the membrane masks and the images of the fluorescent protein (if z-stacks where acquired only the mean-projection of each tile) must be loaded.

→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "4. Quantify Specking Phenotype”
→ Load cell masks and raw phenotype images
→ Click "start anaylsis"

A “spot maximum” for the fluorescent protein within the area of each mask is calculated for each cell and entered in the output file.

6) Combine Data to one Screen file

In the last step all steps of the analysis are combined with the library dictionary to create a final Screen file. This file consists of all cells that have a matching library sequence based on their NIS-Seq signals and a matching phenotype cell based on the alignment of images via the cells coordinates. Additionally, it contains the correlation/spot value for the investigated phenotype. The final file can be analyzed and gated using the visual data exploration tool.

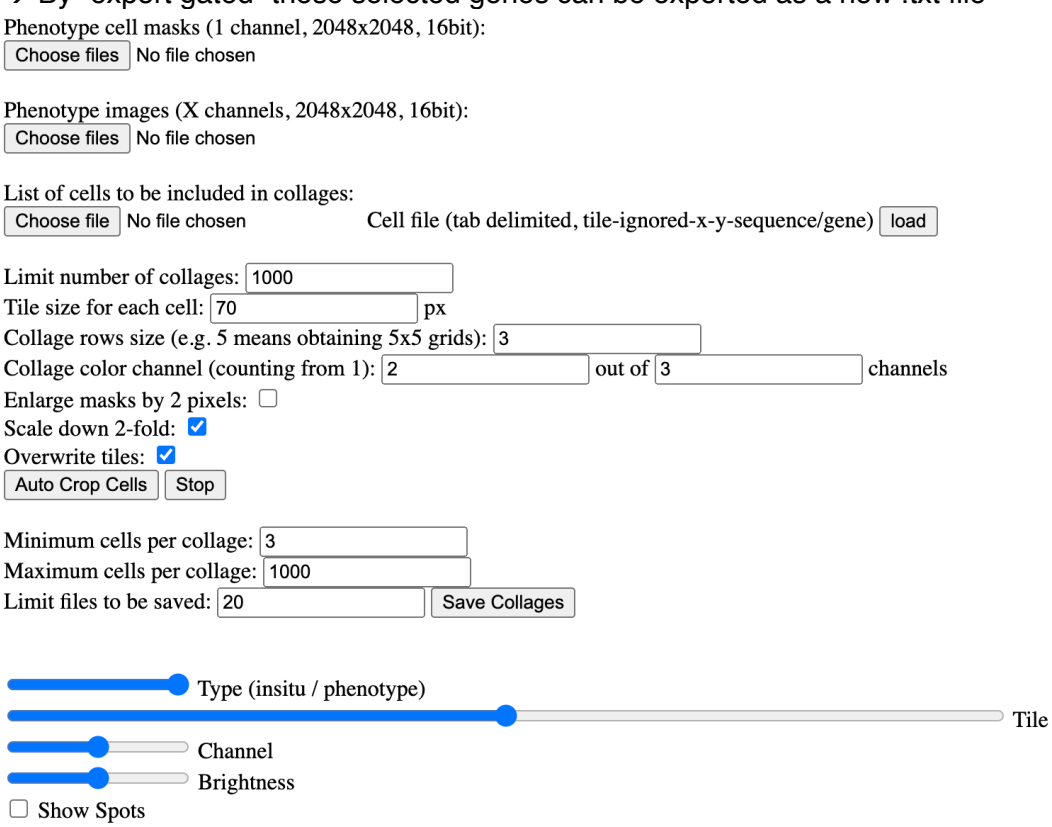


→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "5. Create final Screen file"
→ load the Cropping file created in step 2 at “nuclei”
→ load the assignment file created in step 4 at “nuclei assignment”
→ load the phenotype file created in step 5 at “phenotype”
→ load the dictionary of your knockout library at “guide library”
→ Click “read files” to load all
→ Click “start” to combine all to the final Screen file

7) Obtain Collages

In order to obtain collages from gene specific cells in the screen, the final screen file created in step 6 must be loaded into the visual data exploration tool.

→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click "6. Visual Data Exploration"
→ load the screen file as .txt
→ By changing the X and Y parameter the required parameters can be plotted
→ Click "gate none" to deselect all cells/data points in the plot
→ Use the "search" function to highlight a defined set of genes which will be listed in the window "gated objects"
→ By "export gated" these selected genes can be exported as a new .txt file



This table will serve as the input in the collage tool at "list of cells to be included in collages"
→ In Firefox or Chrome browser, navigate to jsb-lab.bio/opticalscreening/
→ Click " 7. Obtain Collages"
→ load the phenotype cell masks and phenotype images

- if three channels are loaded e.g. nucleus staining, membrane staining and fluorescent reporter the collage color channel must be set to the channel of interest
- if more than one channel is required in the collage the tool has to run once for each channel with the appropriate color channel number; collages of different channels can be merged later in an image analysis tool

→ load the.txt file of gated cells to be included in the collages
→ adjust the parameters to define e.g. number of cells depicted in the collage, the cell size or the channel that is used for the collage generation
→ Click "auto crop cells" (the tool runs through all images and crops the cells given in the list)
→ Click "save collages" to create collages of these cropped cells with the given parameters
→ each collage is saved separately with the name of the gene targeted in the depicted cells