

Explanations on the scripts to generate Multiplexed Protein Maps

Accompanying the publication:

G. Gut et al., Science 361, eaar7042 (2018). DOI: 10.1126/science.aar7042

On directory structure

Script_01_GenerateMPPMatrix.m and Script_05_Proximity_Size_Calculation.m and their subfunctions assume the following directory structure, when looking for images, and segmentations illumination correction bias files. Change your directory structure or parts of the functions, responsible for file system handling. **Importantly, the first directory listed (alphabetically) in the project directory will be treated as Cycle01, the last directory listed in the project directory will be treated as Cycle n.**

Directory structure

Project (directory as specified in "path_project" of Script_01_GenerateMPPMatrix.m)

Cycle01 (directory name must contain string as specified in "str_sharedfoldername" of Script_01_GenerateMPPMatrix.m)

TIFF (directory containing 16-bit images, normally .tiff or .png)

SEGMENTATION (directory containing 16-bit images of segmented objects, e.g. Cells)

BATCH (directory containing illumination bias correction files)

Cycle02 (directory name must contain string as specified in "str_sharedfoldername" of Script_01_GenerateMPPMatrix.m)

TIFF (directory containing 16-bit images, normally .tiff or .png)

SEGMENTATION (directory containing 16-bit images of segmented objects, e.g. Cells)

BATCH (directory containing illumination bias correction files)

Cycle n (directory name must contain string as specified in "str_sharedfoldername" of Script_01_GenerateMPPMatrix.m)

TIFF (directory containing 16-bit images, normally .tiff or .png)

SEGMENTATION (directory containing 16-bit images of segmented objects, e.g. Cells)

BATCH (directory containing illumination bias correction files)

Script_01_GenerateMPPMatrix.m expect a specific name for illumination bias correction file. You might have to change that too. For channel 1 of a given 4i cycles it is

"Measurements_batch_illcor_channel001_zstack000.mat", for channel 2

"Measurements_batch_illcor_channel002_zstack000.mat", etc.

Alternatively, perform illumination bias correction separately from the MPM scripts, and deposit already corrected images into the TIFF folder. If illumination correction bias is neglect able in your images, replace the function “imread_illumination_corrected” by “imread”.

On the file name and channel specification

Scripts to generate Multiplexed Protein Maps expect the following file name and channel specifications for **images (stored in TIFF directory)**.

ImageNameExample_ **A01**_T0001**F001**L01A01Z01**C01**.png

The **yellow marked part** specifies the row and the column of a multiwell plate, in which the image was taken. Specifically, the letter A - N the row and the digits 01 – 24 the columns. The **green marked part** specifies the microscopy site within a well. E.g. digits may go from F001-F024. The **cyan marked part** specifies the channel number. Specifically, the letter C will not change, however, the subsequent digit may. E.g. C01 for the first channel acquired (DAPI), C02 for the second channel acquired (AF488), C03 for the third channel acquired (AF568).

Scripts to generate Multiplexed Protein Maps expect the following file name and channel specifications for **segmentation files (stored in TIFF directory)**.

ImageNameExample_ **A01**_T0001**F001**L01A01Z01**C01**_Segmented**Cells**.png

The **yellow**, **green**, and **cyan** parts reflect information as introduced above. The **magenta marked part** specifies the type of segmentation object stored in the segmentation file, e.g. Cells, Cytoplasm or Nuclei.

Image analysis was performed using CellProfiler1, thus output of segmentation modules is according to CP1 standards. E.g. 16-bit image where the segmentation is uniformly marked with a number, which corresponds to the number in the 4th columns expected in the cell metadata file (see below).

On the cell metadata

Scripts to generate Multiplexed Protein Maps require a CSV file which contains metadata information about the location and ID of the cells from which the multiplexed pixel profiles should be extracted. Rows of the CSV file correspond to cells. The first column specifies the row of the well in which the cell is located (A = 1. B = 2, C = 3, etc.). The second column specifies the columns of the well in which the cell is located. The third columns specifies in which **Absolut** microscopy site a cell is located. Assuming an experiment has 3 wells, C03, C04, C05, and C6, and in each well 42 sites are imaged. Absolute microscopy site means in this case that a cell located in the well C05 in the first microscopy site will have an absolute microscopy site of 85 (42 sites in C03 + 42 sites in C04 + 1 site in C05). The fourth column specifies the Cell ID as measured by CellProfiler, when segmentation is computed (see section above).

Exclude this column from CSV file	Row	Column	Absolute site	Cell ID
Cell 1	3	5	85	25
Cell 2	3	6	146	12