Standard Operating Procedure

# SOP details

Title: Image Preprocessing and Classification Pipeline for TopoChip Analysis

Description: This SOP outlines the steps to preprocess raw TopoChip images, crop regions of interest, run the CellProfiler pipeline, and classify surface responses.

Author: Nikita Kosnhin; Koen Minartz

SOP number: 01

Version number: 1

# 1. Purpose

To prepare aligned single cell on topography image data for modeling by preprocessing, cropping, and classifying regions of interest using MATLAB, Python and CellProfiler. This includes removing bad tiles, scoring quality, and exporting data for downstream machine learning.

# 2. Principle

This SOP describes the step-by-step procedure to:  
- Preprocess raw topography insert images  
- Manually exclude bad tiles  
- Crop images for CellProfiler  
- Run and configure the CellProfiler pipeline  
- Score image quality with CellProfiler Analyst  
- Export and merge quality data  
- Apply fixed image cropping for downstream analyses

# 3. Important to Know Before Starting

1. This SOP is highly dependent on consistent file structure and naming conventions.  
2. Image artifacts and border tiles must be excluded manually before proceeding.  
3. Ensure that input and output folders are emptied before each new batch.  
4. Some steps require cautious manual selection of representative and high-quality image tiles.  
5. Fixed image regions should closely match the best-quality area for fair downstream comparison.

# 4. Required Materials

4.1 Workplace  
Desktop or laptop with installed:  
- MATLAB  
- CellProfiler  
- CellProfiler Analyst  
- Jupyter Notebook (Python environment)  
  
4.2 Requirements  
- Access to “main\_Koen.m” MATLAB script  
- CellProfiler pipeline and classifier templates  
- Raw image data  
- "Phenome28.doc" for metadata consistency

## 5.1 Preprocessing and Tile Selection

1. Clear Input and Output folders using the provided Jupyter notebook.

2. Copy raw data into the RawData folder.

3. Ensure naming consistency using Table 1 as reference.

Table 1. Raw data naming structure

|  |  |
| --- | --- |
| Filename | Channel Explanation |
| ChannelTRITC-1,FITC3,DIA-Ph3,DIA-Ph4,DAPI2\_Seq0000\_DAPI2.tif | DAPI2 → nuclei staining |
| ChannelTRITC-1,FITC3,DIA-Ph3,DIA-Ph4,DAPI2\_Seq0000\_DIA-Ph3.tif | DIA-Ph3 → Brightfield |
| ChannelTRITC-1,FITC3,DIA-Ph3,DIA-Ph4,DAPI2\_Seq0000\_FITC3.tif | FITC3 → Actin staining |
| ChannelTRITC-1,FITC3,DIA-Ph3,DIA-Ph4,DAPI2\_Seq0000\_TRITC-1.tif | TRITC-1 → YAP staining |

4. Launch MATLAB and open main\_Koen.m.

5. Set parameters: run\_SelectTilesForCP = 1, all others = 0. Run the script.

6. Visually inspect tiles. At line 85, mark and exclude: border tiles, overlapping chip borders, and artifacts.

7. Record best tile numbers for step 16.

## 5.2 Cropping and Validation

8. Ensure Input/CP\_input\_images is empty.

9. Re-run main\_Koen.m with run\_CropCPImages = 1 and run\_SelectTilesForCP = 1.

10. Validate cropped tiles by visual inspection.

11. Cross-check tile IDs and topounit codes with Phenome28.doc.

## 5.3 Running CellProfiler

12. Start the test mode, test 20 random images for the segmentation performance, adjust IdentifyPrimaryObjects and IdentifySecondaryObjects if needed

13. Open CellProfiler pipeline and configure ExportToDatabase with:  
 - ActinWithoutNeighbours  
 - NucleiNoNeighbours  
 - Set location to NucleiNoNeighbours

14. Save pipeline in cellprofiler\_protocols\_for\_all\_the\_chips/[CHIPNUMBER]/

15. Save input/output in batch/[chip name]/CP\_input and CP\_output folders.

## 5.4 Image Quality Classification

16. Launch CellProfiler Analyst → classifier:  
 - Score images  
 - Save model as QC\_classifier\_model\_Feature\_Idx\_...  
 - Save table as QC\_classifier\_table.csv (not MyExpt\_per\_Image)

## 5.5 Merge and Filter QC Output

17. Use Concat\_classifier\_output\_with\_tables.ipynb to:  
 - Merge NoNeighboursNucleiNoNeighbours.csv with QC\_classifier\_table.csv  
 - Remove negative objects  
 - Repeat for ActinWithoutNeighbours

## 5.6 Outlier Removal

18. Re-run main\_Koen.m with run\_outlierRemoval = 1 and others = 0.

## 5.7 Fixed Image Cropping

19. Select best-looking tile:  
 - Set run\_SelectTilesForCP = 1, others = 0  
 - Choose tile from Input/CP\_input\_images

20. At line 106 in main\_Koen.m, enter coordinates for fixed tile.

21. Set run\_FixedImagesCropper = 1 and run.  
 - Ensure region is optimal.  
 - Copy result to cellprofiler\_protocols\_for\_all\_the\_chips/