

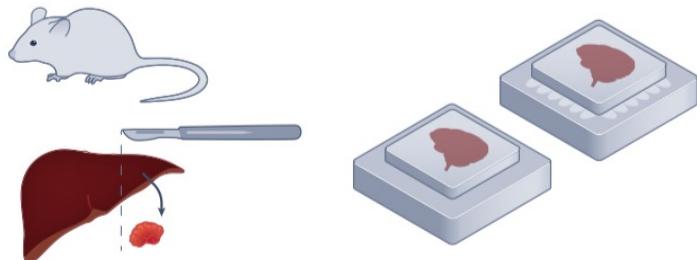
MCMICRO & SCIMAP Workshop

2 June 2023

By Kriengkrai Phongkitkarun (Heng)

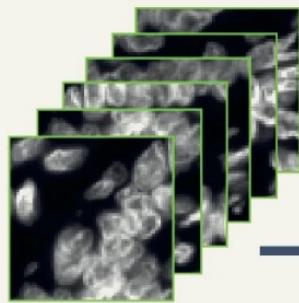
1. Biospecimen metadata level

Tissue biopsy, paraffin embedding, tissue sectioning and antigen retrieval

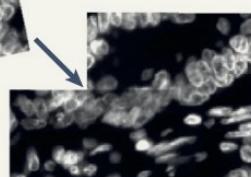


MCMICRO

a MITI data levels 1 and 2 (access may be limited):



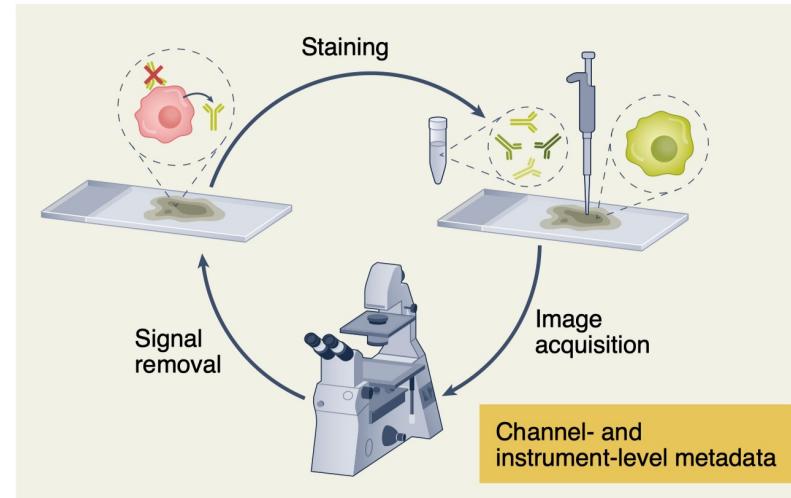
Biospecimen metadata



Level 1 data:
Raw image tiles
.tiff, .svs, .ims, .czi, .mcd,
.rcpnl, .dicom, .nd2, .xdce, etc.

Level 2 data:
Assembled, multi-
channel images
.ome.tif

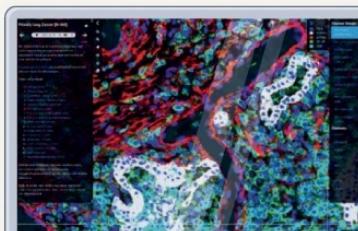
2. Experimental metadata level



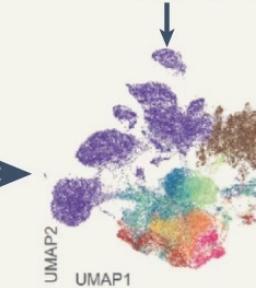
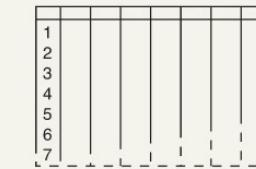
Staining proteins/molecules
Imaging immunofluorescent

SCIMAP

b MITI data levels 3–5 (unrestricted access):

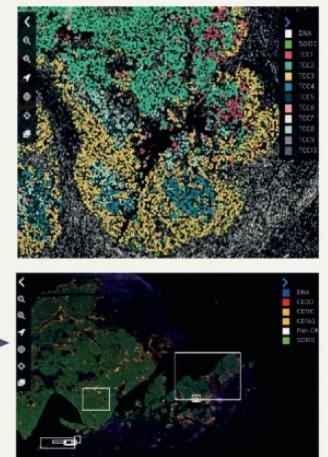


Level 3 data:
Quality controlled, assembled
images, segmentation masks
.ome.tif



Level 4 data:
Spatial feature table
.fcs, .csv, .h5ad

3. Data level



Level 5 data:
Data models
Annotated images
MINERVA stories

Four computational challenges in multiplexed tissue imaging:

- (1) data acquired in multiple image fields must be **assembled precisely** into large mosaic images in the whole specimen and multiple imaging cycles
- (2) **full-resolution images** must be made available in conjunction with **numerical results**
- (3) images must be **segmented into single cells** - a difficult task when cells are densely crowded and nuclei have irregular morphologies
- (4) diverse image-processing algorithms and data types must be **harmonized** across research groups and programming languages



OPEN

MCMICRO: a scalable, modular image-processing pipeline for multiplexed tissue imaging

Denis Schapiro^{1,2,3,18,20}, Artem Sokolov^{1,2,4,20}, Clarence Yapp^{1,2,5,20}, Yu-An Chen^{1,2},
Jeremy L. Muhlich^{1,2}, Joshua Hess^{1,6}, Allison L. Creason^{1,7}, Ajit J. Nirmal^{1,2,8},
Gregory J. Baker^{1,2}, Maulik K. Nariya^{1,2}, Jia-Ren Lin^{1,2}, Zoltan Maliga^{1,2}, Connor A. Jacobson^{1,2},
Matthew W. Hodgman^{2,9}, Juha Ruokonen^{1,2}, Samouil L. Farhi^{1,3}, Domenic Abbondanza³,
Eliot T. McKinley^{10,11}, Daniel Persson^{7,12}, Courtney Betts¹³, Shamilene Sivagnanam¹³, Aviv Regev^{1,3,14,19},
Jeremy Goecks^{1,7,12}, Robert J. Coffey^{1,15}, Lisa M. Coussens^{1,12,13}, Sandro Santagata^{1,2,16} and
Peter K. Sorger^{1,2,17}✉

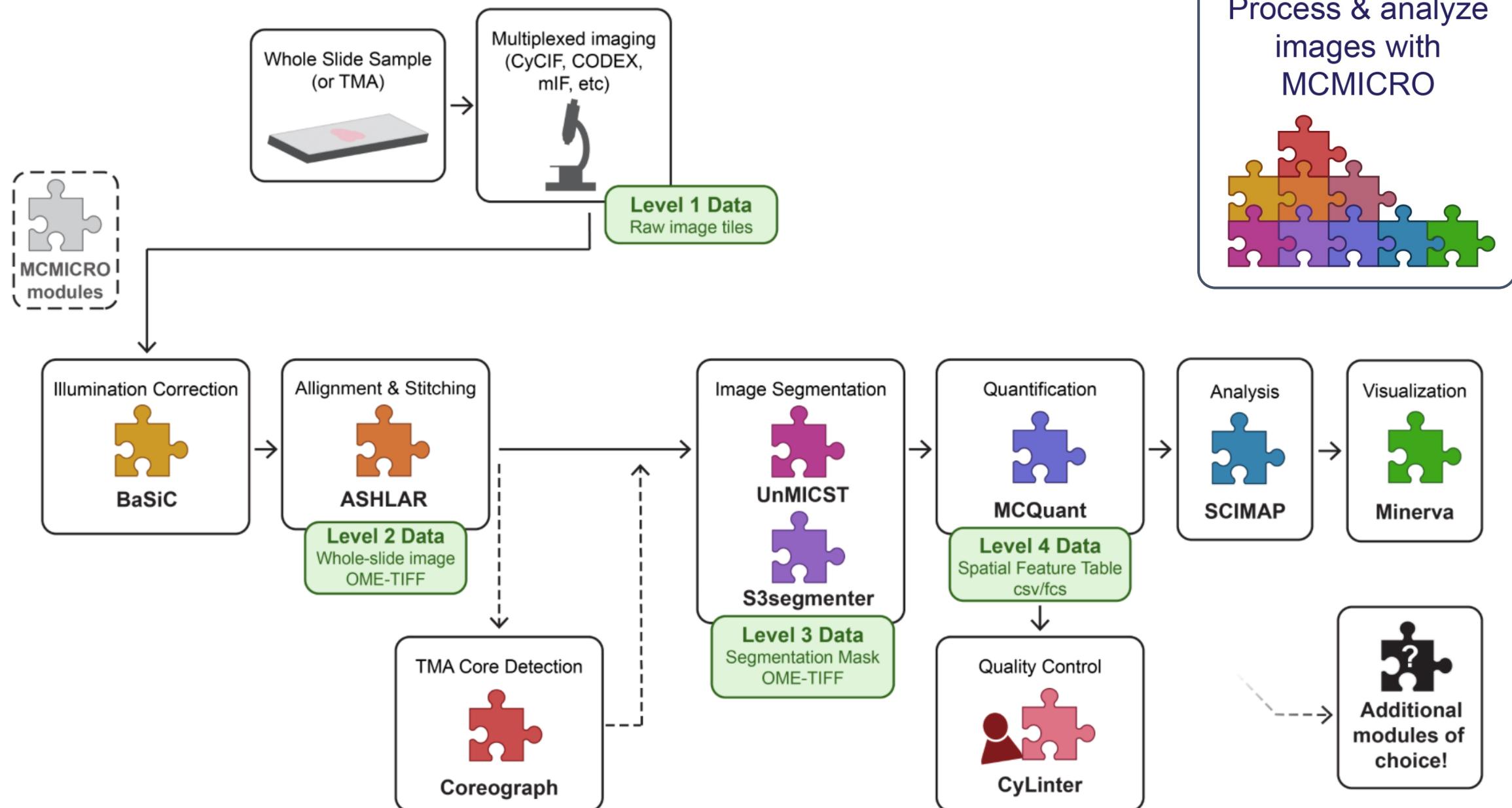
Multiple-choice microscopy pipeline

An end-to-end processing pipeline that transforms multi-channel whole-slide images into single-cell data. This website is a consolidated source of information for when, why, and how to use MCMICRO.

[OVERVIEW →](#)[TUTORIAL →](#)[GITHUB →](#)

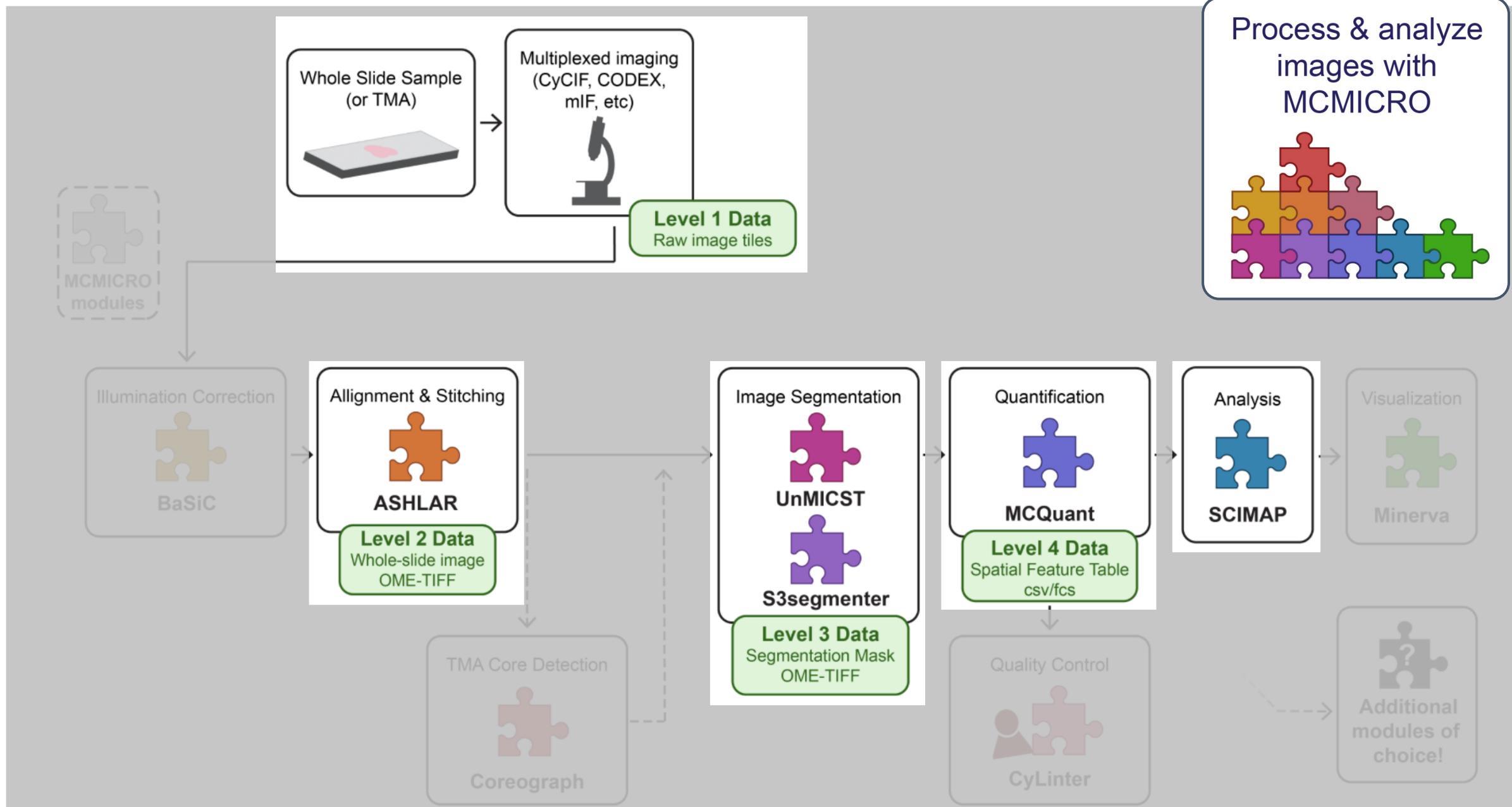
MCMICRO, a **modular** and open-source computational pipeline, for performing the sequential steps needed to transform whole-slide images into single-cell data

»» <https://mcmicro.org>



MCMICRO, a **modular** and open-source computational pipeline, for performing the sequential steps needed to transform whole-slide images into single-cell data

»» <https://mcmicro.org>



Preparation files/programs for using MCMICRO & SCIMAP

Files

- | | |
|--------------------------------------|------------------------------------|
| 1. Raw unstitched images | .rcpnl from RareCyte |
| 2. <code>markers.csv</code> | image metadata (cycles & channels) |
| 3. <code>param.yml</code> | parameter files for MCMICRO |
| 4. <code>phenotype.csv</code> | user-defined cell phenotypes |

Programs/Installation

- | | |
|----------------------------|---|
| 1. Jupyter notebook | Python |
| 2. Google Chrome | to access Jupyter |
| 3. QuPath | image exploration (https://qupath.github.io) |
| 4. Microsoft Excel | table data exploration |

markers.csv

image metadata (cycles & channels)

channel_number	cycle_number	marker_name	filter	excitation_wavelength	excitation_wavelength	exposure(sec)
1	1	None_1	A488	-	-	-
2	1	CD11c	A647	651	692	0.07
3	1	DNA_1	BV421	395	438	0.006
4	1	Ki-67	Sytox	555	590	0.1
5	2	PD-L1	A488	485	522	0.35
6	2	PD-1	A647	651	692	0.25
7	2	DNA_2	BV421	395	438	0.006
8	2	CD3d	Sytox	555	590	0.2
9	3	CD8a	A488	485	522	0.15
10	3	None_2	A647	-	-	-
11	3	DNA_3	BV421	395	438	0.006
12	3	CD68	Sytox	555	590	0.25
13	4	CD4	A488	485	522	0.1
14	4	PD-1	A647	651	692	0.02
15	4	DNA_4	BV421	395	438	0.006
16	4	FOXP3	Sytox	555	590	0.15
17	5	CD163	A488	485	522	0.1
18	5	SMA	A647	651	692	0.15
19	5	DNA_5	BV421	395	438	0.006
20	5	PanCK	Sytox	555	590	0.15

phenotype.csv

user-defined cell phenotypes

	CD11c	Ki-67	PD-L1	PD-1	CD3d	CD8a	CD68	CD4	CD20	FOXP3	CD163	SMA	PanCK
proliferating cells		pos											
dendritic cells	pos												
exhausted T cell			pos	pos	pos								
Treg						pos			pos		pos		
cytotoxic T cell						pos	pos						
helper T cell						pos	pos						
M1 macrophage								pos				neg	
M2 macrophage								pos			pos		
epithelial cell													pos
stroma													
immunoregulatory protein		pos	pos										
B cell									pos				

param.yml

parameter files for MCMICRO

```
workflow:
  start-at: registration
  stop-at: downstream
  qc-files: copy
  tma: false
  viz: true
  background: false
  multi-formats: '{.xdce,.nd,.scan,.htd}'
  single-formats: '{.ome.tiff,.ome.tif,.rcpnl,.btif,.nd2,.tif,.czi}'
  segmentation: unmict
  segmentation-recycle: false
  downstream: scimap
options:
  ashlar: -m 30
  cypository: --model zeisscyto
  ilastik: --num_channels 1
  mcquant: --masks cell*.tif
  naivestates: -p png
modules:
  illumination:
    name: basic
    container: labsyspharm/basic-illumination
    version: 1.1.1
  registration:
    name: ashlar
    container: labsyspharm/ashlar
    version: 1.17.0
  dearay:
    name: coreograph
    container: labsyspharm/unetcoreograph
```

```
name: cypository
container: labsyspharm/cypository
version: 1.1.5
cmd: python /app/deployMaskRCNN.py --stackOutput --outputPath .
input: "
channel: --channel
idxbase: 1
watershed: bypass
-
name: ilastik
container: labsyspharm/mcmicro-ilastik
version: 1.6.1
cmd: python /app/mc-ilastik.py --output .
input: --input
model: --model
channel: --channelIDs
idxbase: 1
watershed: 'yes'
-
name: mesmer
container: vanvalenlab/deepcell-applications
version: 0.4.0
cmd: python /usr/src/app/run_app.py mesmer --squeeze --output-directory . --output-name
cell.tif
input: --nuclear-image
channel: --nuclear-channel
idxbase: 0
watershed: 'no'
-
```

Multiplexed tissue imaging workflow in SiSP

PURPOSE/Input

1 Generating MCMICRO input

- Unstitched raw imaged in Bioformat
- .rcpnl from RareCyte

2 Image illumination correction

- Background shading correction
- Producing high accuracy
- Only require few images (4 fields)

3 Image Alignment & Stitching

- Image registration
- By nuclear channels between cycles
- Required non-saturated, sharp images

4 Image segmentation mask

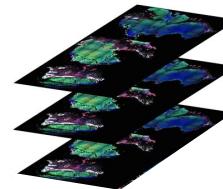
- Nuclear, cell segmentation each markers

5 Spatial feature quantification

- Single-cell level data
- Quantify intensity markers, morphology (size, shape, granularity...), coordinates (X,Y)

PROCESS

Cycle1
Cycle 2
Cycle 3



t-cycif, whole slide imaging

BaSic



ASHLAR



UnMICST

*llastik



MCQuant



TOOLS/Output

RareCyte

.rcpnl

QuPath

Registered images

- Check registered output images
- Check intensity markers, sharpness

QuPath

Segmented images

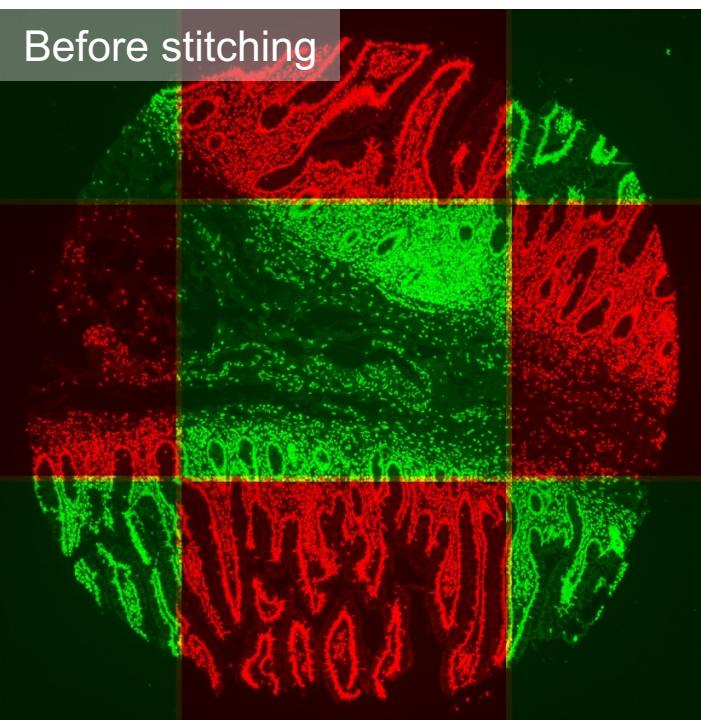
- Check segmentation correction
- If found wrong segmentation, considering adjust/change the algorithms

Table

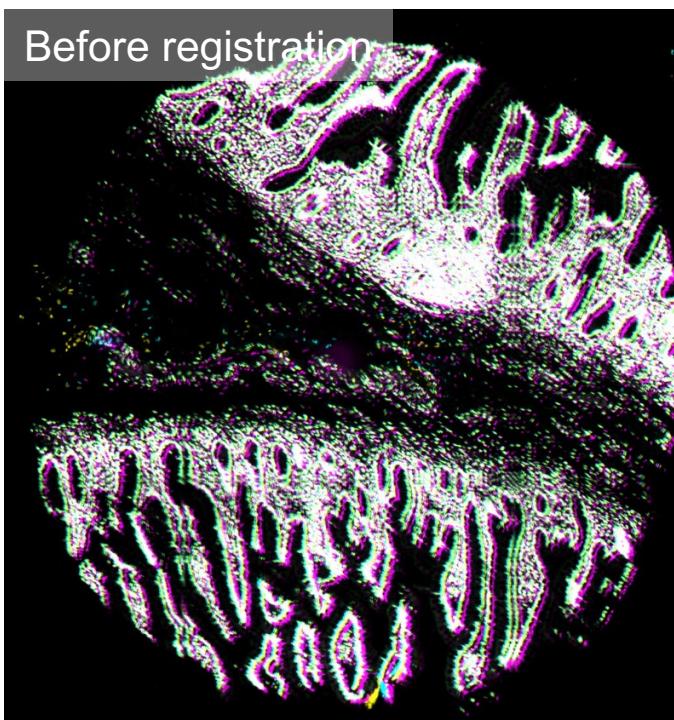
.csv

- Check row table, representing single-cell level data

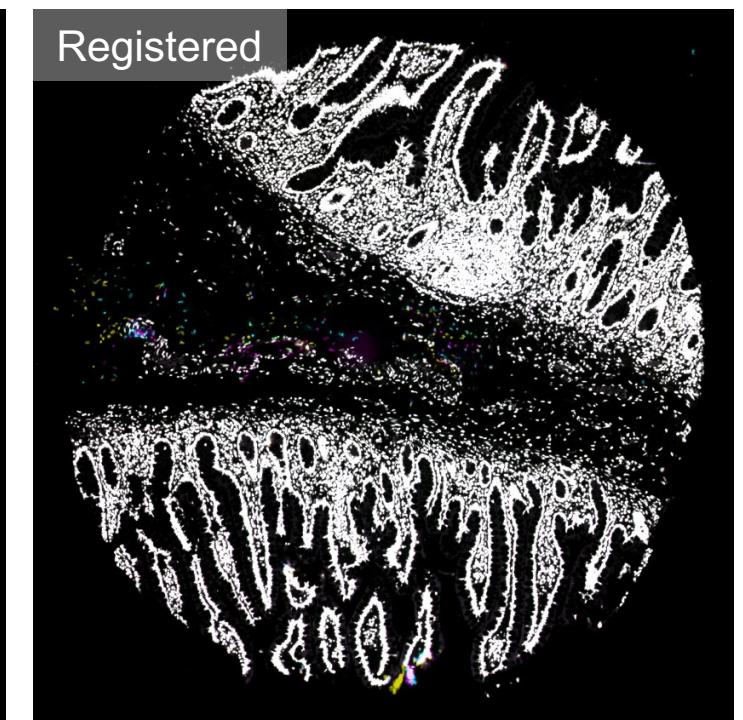
Before stitching



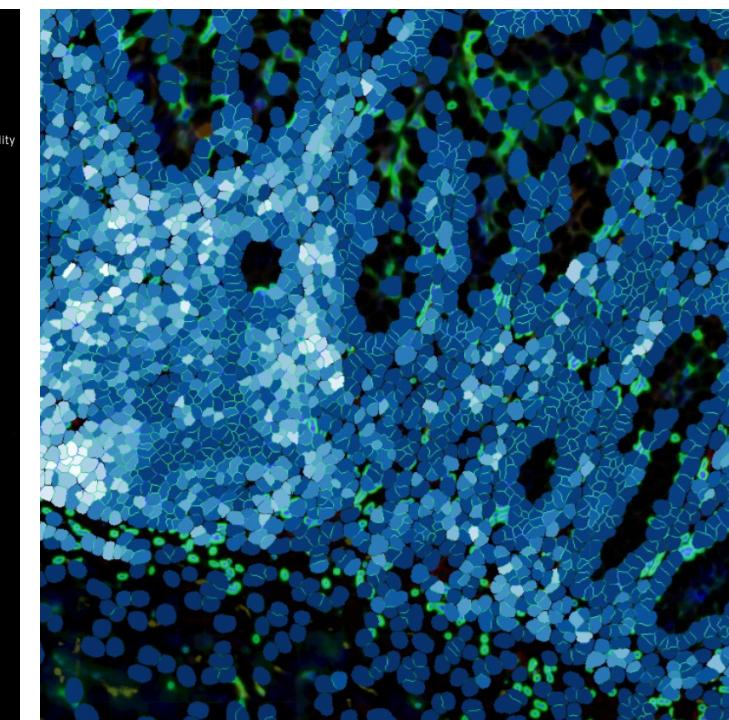
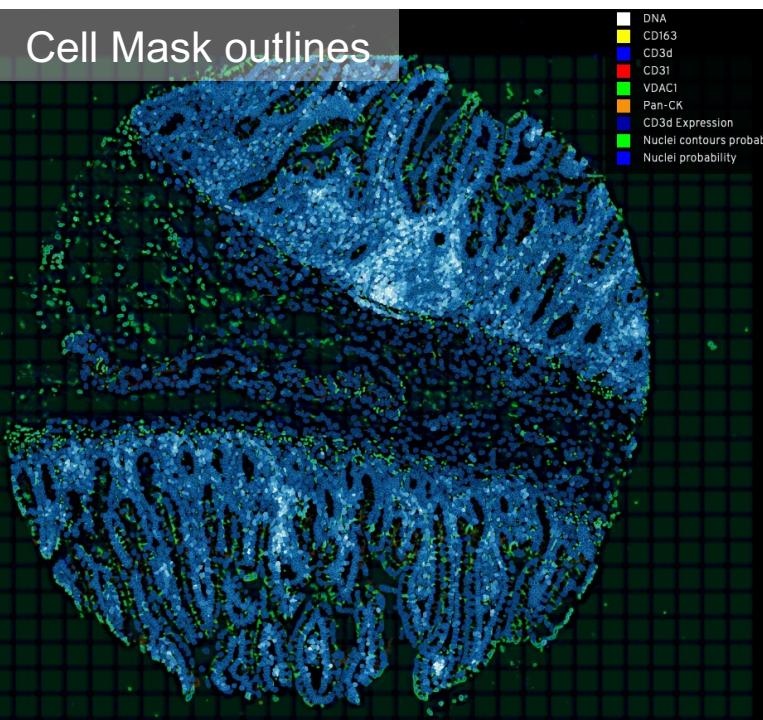
Before registration



Registered



Cell Mask outlines



<https://mcmicro.org/overview/pipeline-visual-guide>

Multiplexed tissue imaging workflow in SiSP

PURPOSE/Input

6 Generating analysis & visualization

► Required .csv from 5 & phenotype.csv

1) Marker intensity gating

To define intensity marker gating each channel

2) Phenotype mapping

To visual spatial phenotypes

PROCESS



*use stand-alone

- Manual gating
- GMM gating** (Intensity Rescale)

4 algorithms for phenotype clustering

- User defined phenotypes
- Kmeans
- Leiden
- Phenograph

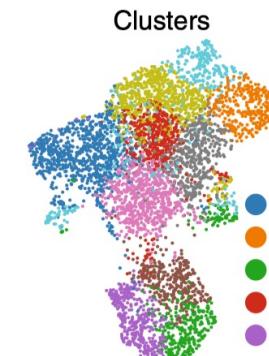
3 visualizing plots

- UMAP clustering
- Neighborhoods
- Cell-type assignment

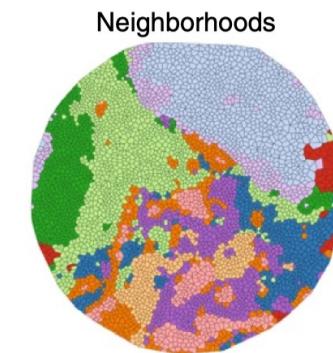
GMM gating

imageid	exemplar-001--unmicst_cell
markers	
CD11B	3.559838
CD16	6.272904
CD45	6.512384
CD57	6.516099
ECAD	3.685633
ELANE	7.392729
FOXP3	6.174185
NCAM	6.352898
SMA	5.538066

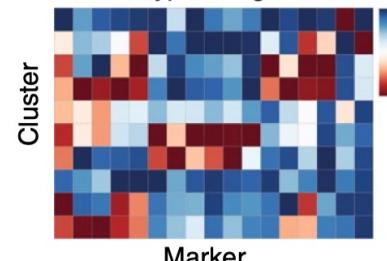
Clusters



Neighborhoods



Cell-type assignment

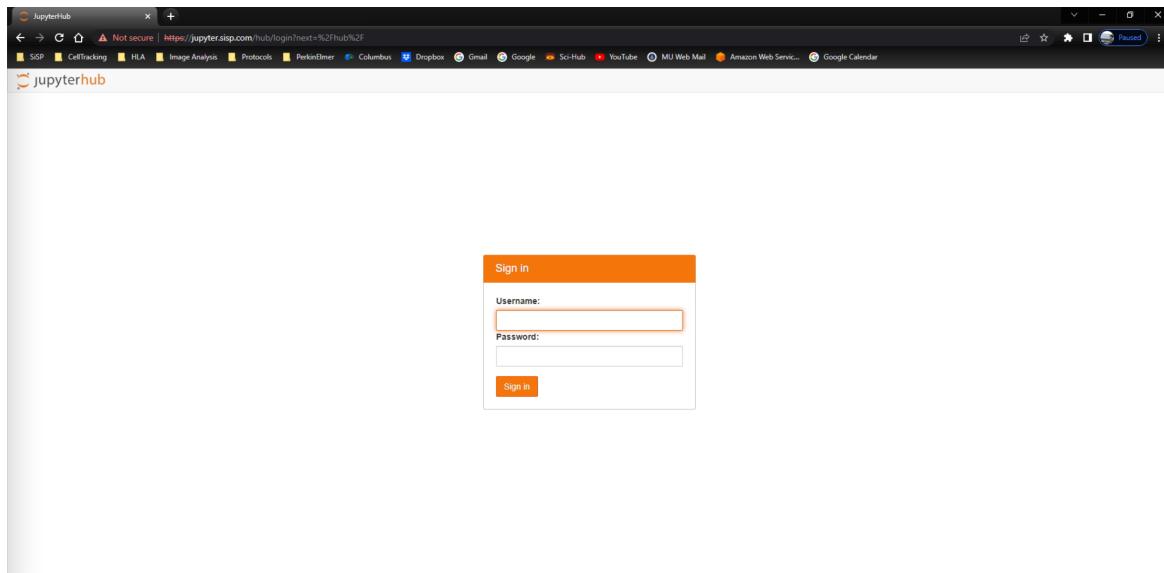


THANK YOU

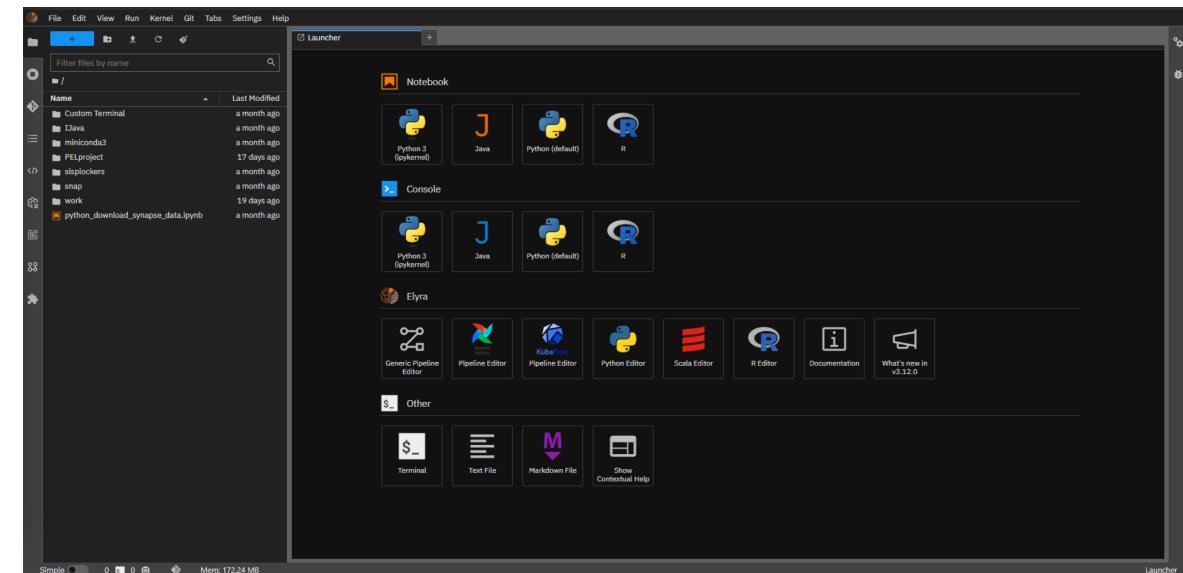
JupyterLab

URL >> <https://jupyter.sisp.com>

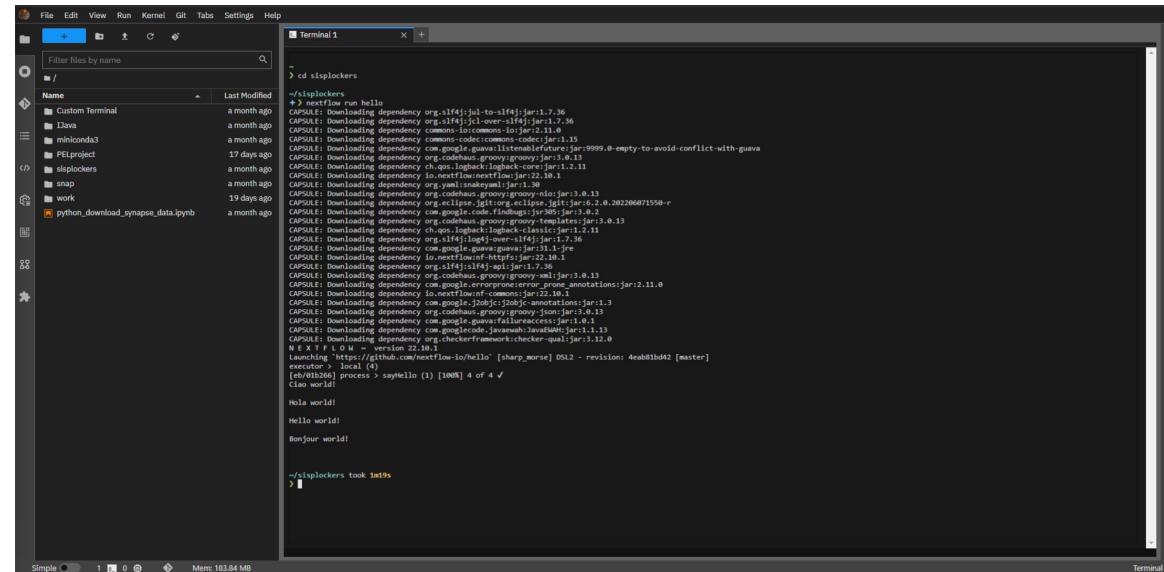
1.



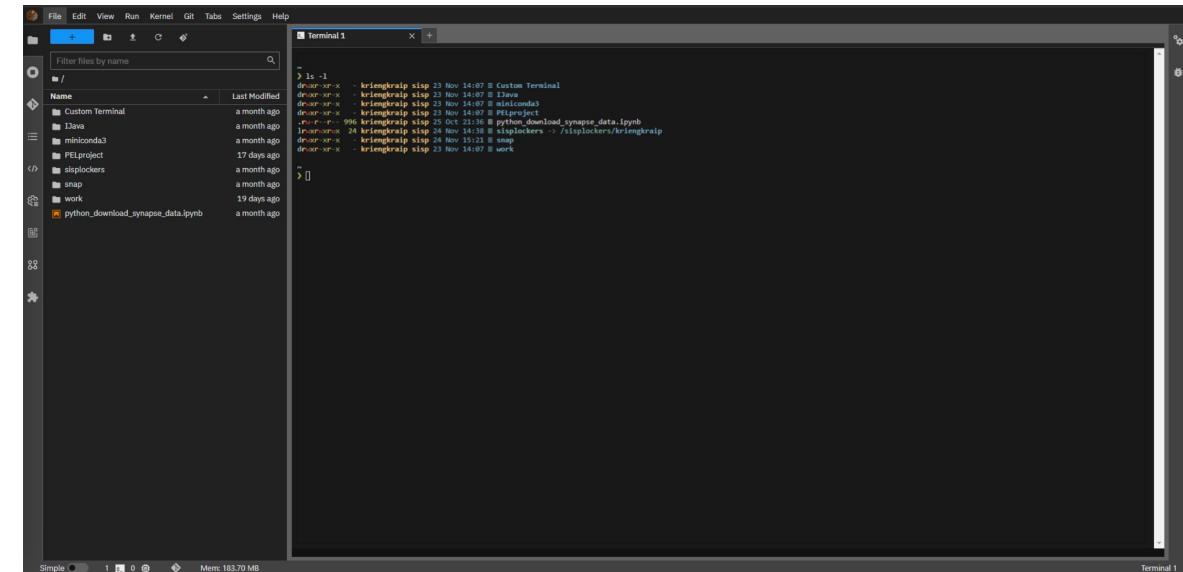
2.



3.



4.

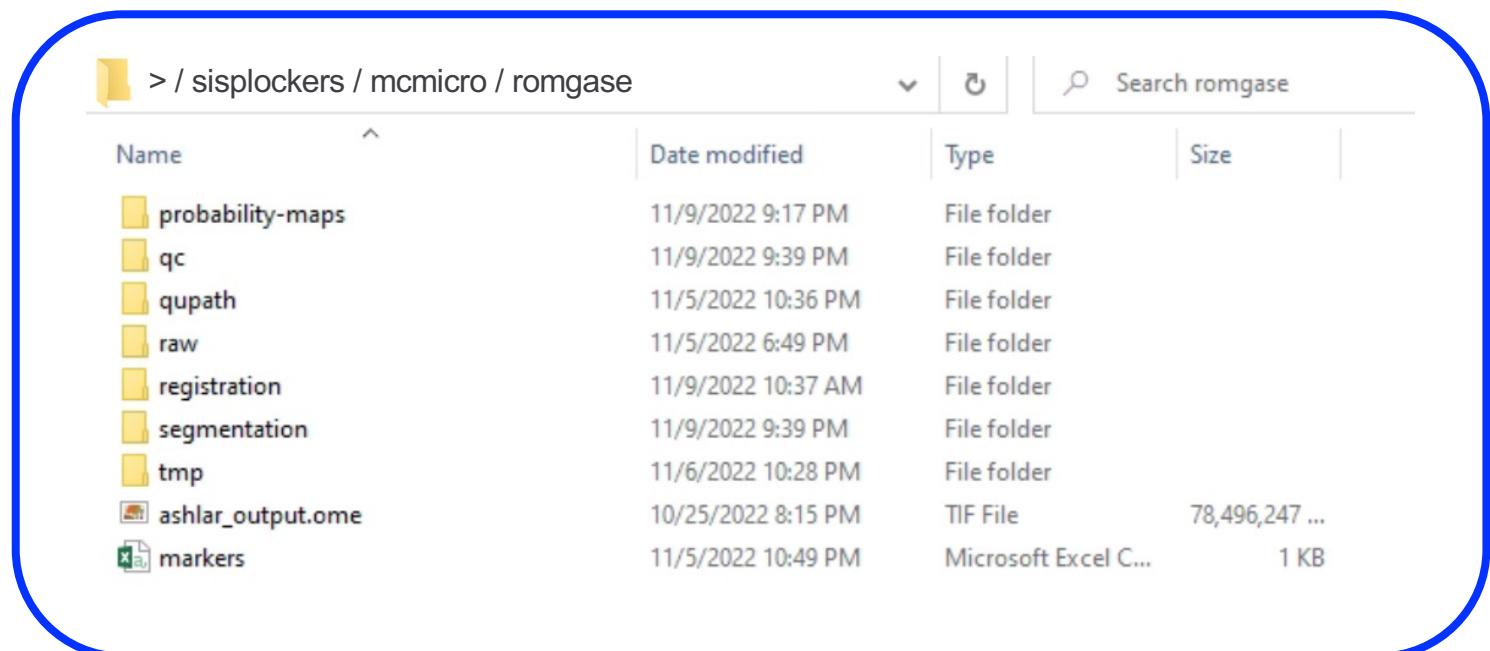
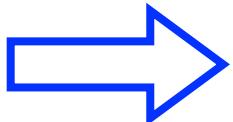


Step to run MCMICRO Commands in JupyterLab

1. Open Google Chrome browser
2. Go to address > <https://jupyter.sisp.com>
3. Login using username and password
4. Click Terminal application
5. Type the following command and press 'Enter'
`$ nextflow run labsyspharm/mcmicro --in /sisplockers/mcmicro/romgase -w ~/work`
6. Wait until finished (~3 hours)

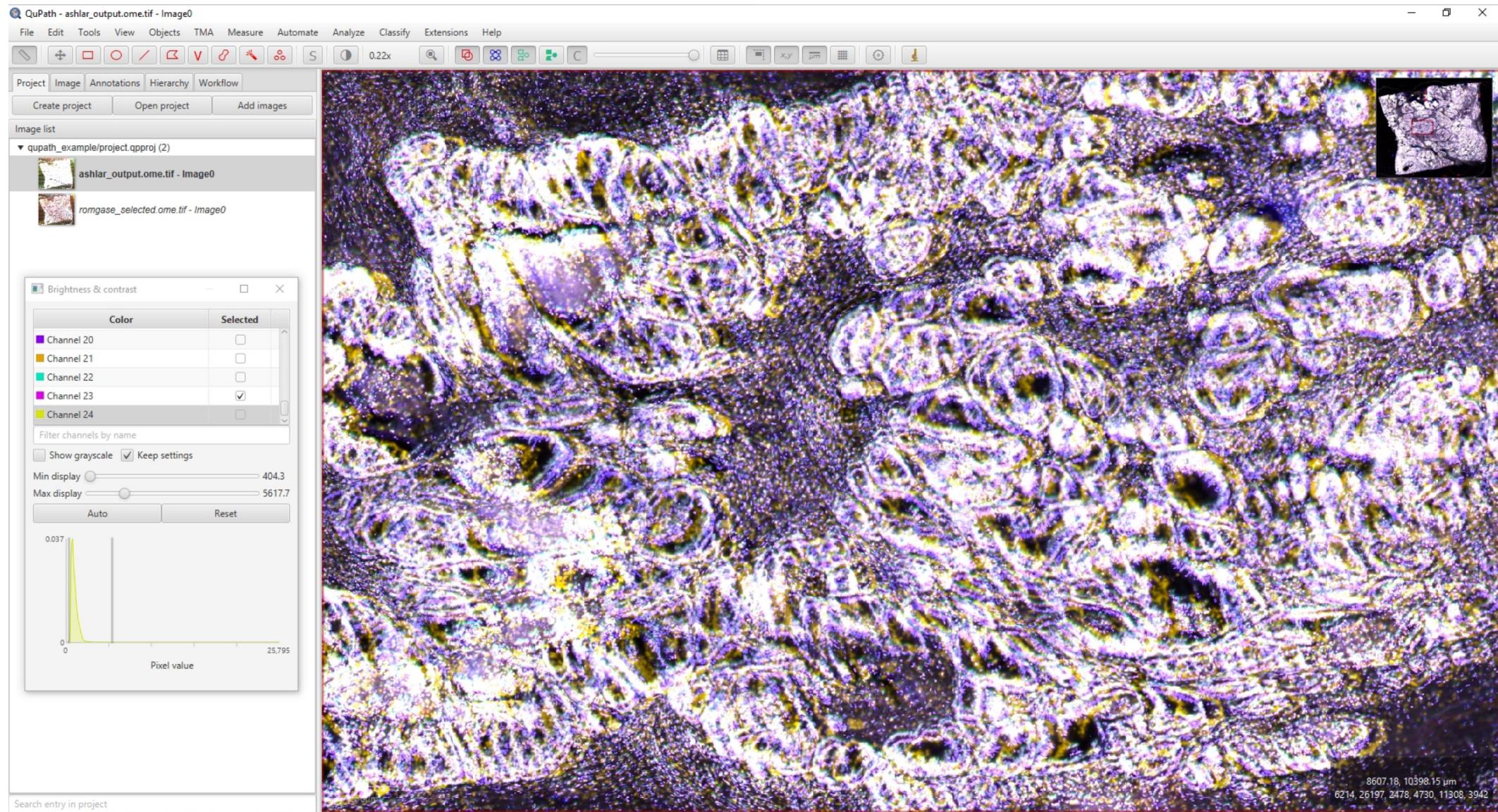
*** This test was run by SISP computer.

Output Files



Name	Date modified	Type	Size
probability-maps	11/9/2022 9:17 PM	File folder	
qc	11/9/2022 9:39 PM	File folder	
qupath	11/5/2022 10:36 PM	File folder	
raw	11/5/2022 6:49 PM	File folder	
registration	11/9/2022 10:37 AM	File folder	
segmentation	11/9/2022 9:39 PM	File folder	
tmp	11/6/2022 10:28 PM	File folder	
ashlar_output.ome	10/25/2022 8:15 PM	TIF File	78,496,247 ...
markers	11/5/2022 10:49 PM	Microsoft Excel C...	1 KB

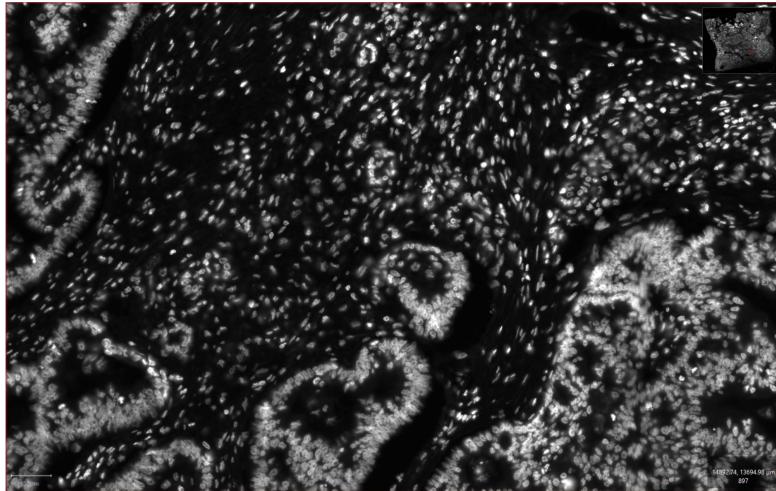
An Alignment Result (from 6 cycles)



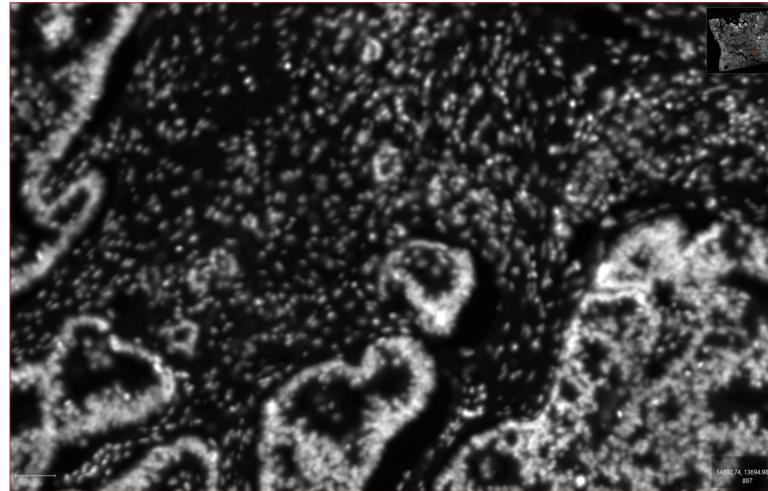
Raw images (from 6 cycles)

Some cycles of the Nuclei channel have blurred images

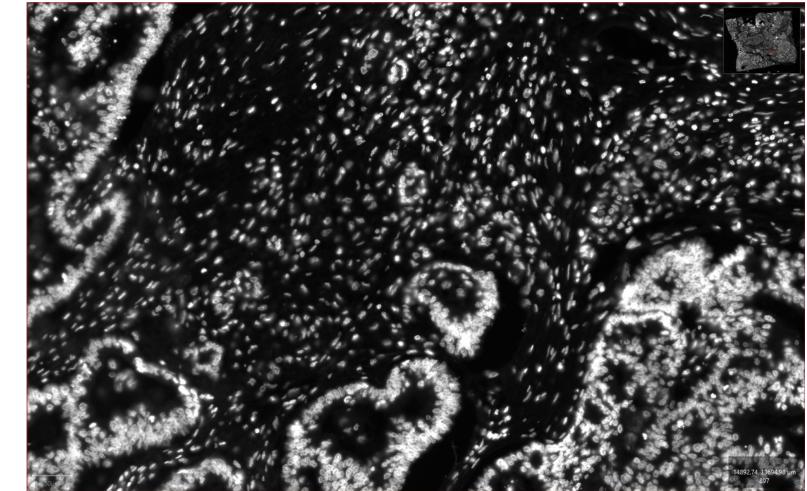
Cycle 1



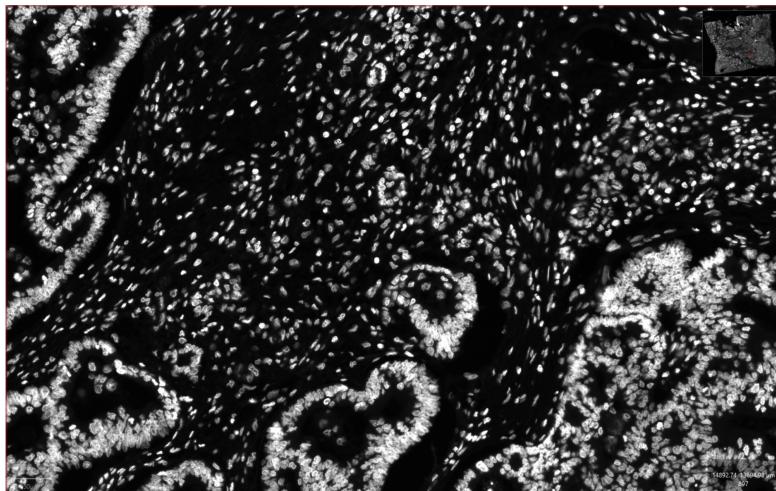
Cycle 2 X



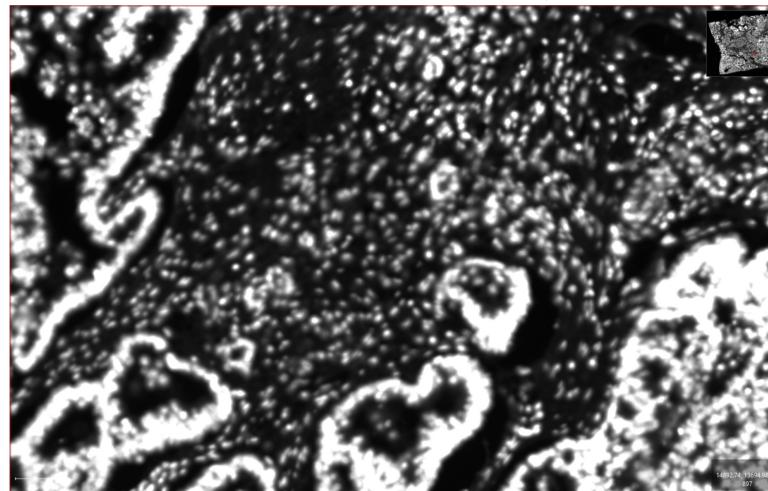
Cycle 3



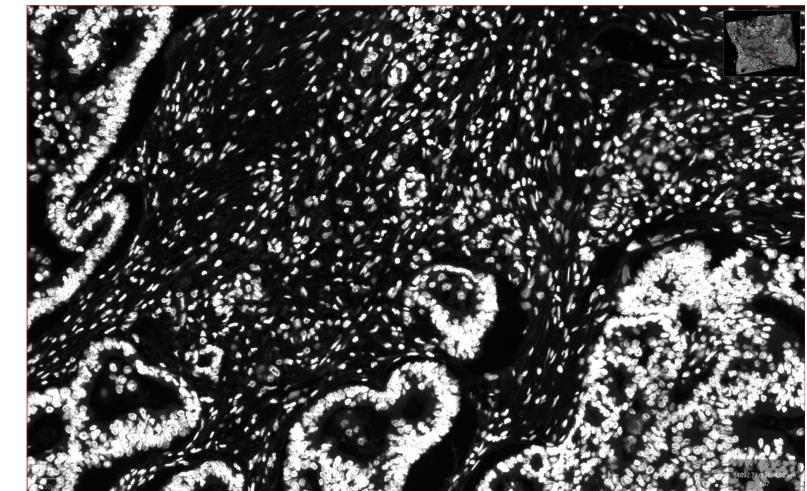
Cycle 4



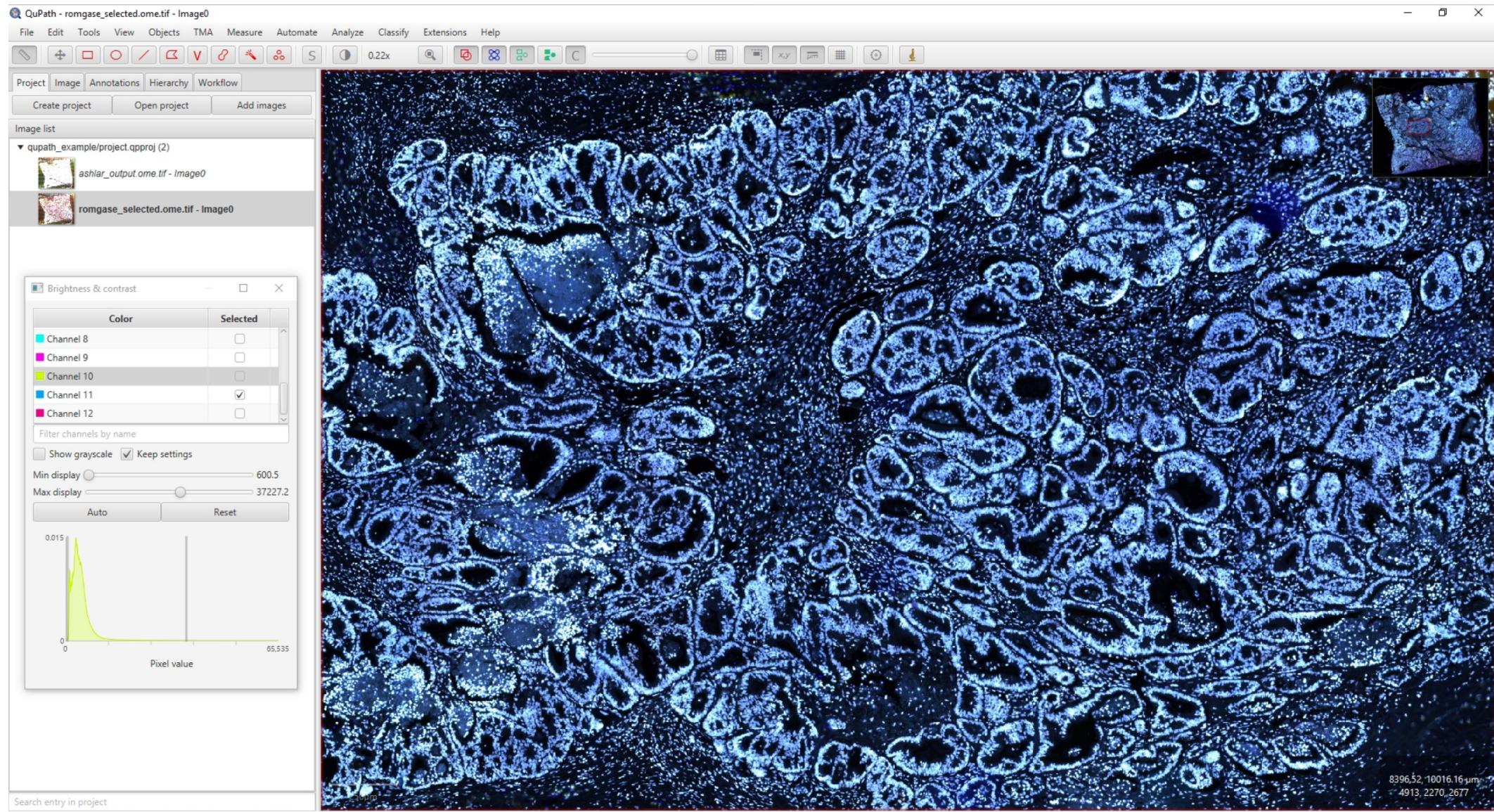
Cycle 5 X



Cycle 6

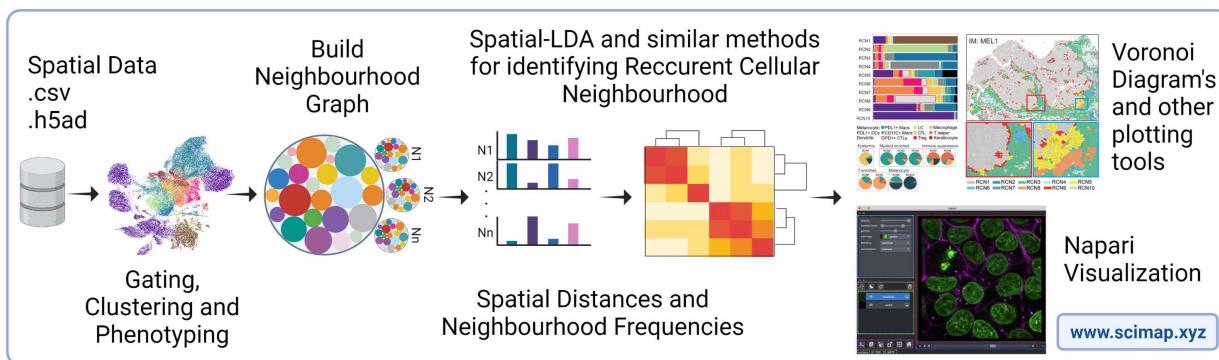


Result images (New version with excluding Cycle 2 and 5)



Scimap in JupyterLab (use Python3 Notebook)

URL >> <https://jupyter.sisp.com>



“Scalable toolkit for analyzing spatial molecular data”

The screenshot shows the Scimap JupyterLab interface with the following elements:

- Header:** scimap, GitHub 0.22.0, 30, 8.
- Navigation:** SCIMAP, Getting Started, Tools Shortcut, All Functions, Tutorials.
- Tutorials:** Getting Started with Scimap, Cell-Phenotyping using Scimap, 3 Cell Type calling and adding ROIs, 4 CellType Proportion Exploration, 5 Simple Spatial Analysis, 6 animate with scimap, Releases.
- Code cell:** A code cell containing Python code for creating a file named "test.txt".
- Output cell:** The output of the code cell showing the creation of the file and its contents.
- Table of contents:** Tutorial material, Tutorial video, Load data using AnnData, Load data using scimap's helper function, We can use scanpy package to explore the data.

Getting Started with Scimap

```
1 #!/usr/bin/env python3
2 # -*- coding: utf-8 -*-
3 """
4 Created on Fri Jun 26 23:11:32 2020
5 @author: Ajit Johnson Nirmal
6 Scimap Getting Started tutorial
7 """
```

'\nCreated on Fri Jun 26 23:11:32 2020\n@author: Ajit Johnson Nirmal\nScimap Getting

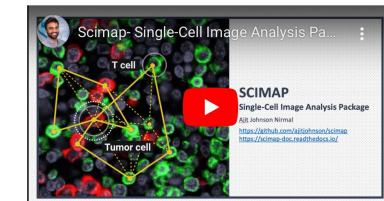
```
1 # Before you start make sure you have installed the following packages
2 # pip install scimap
3 # pip install scanpy
4 # pip install leidenalg
5 # pip install PyQt5
```

Tutorial material

You can download the material for this tutorial from the following link:
The presentation files are available [here](#):

Tutorial video

```
1 from IPython.display import HTML
2 HTML('<iframe width="450" height="250" src="https://www.youtube.com/embed/knh5el...>
```



Ref. <https://scimap.xyz/tutorials/1-scimap-tutorial-getting-started/>