Hyperion data analysis v7

Essential software

- MCD-viewer (any version)
  Used to view MCD files, file conversion and thresholding.
  Software is only available for windows and administrator rights are needed to open.
  Can be downloaded from www.fluidigm.com/software

- Matlab (any version)
  Used for file conversion and image enhancement
  Can be downloaded from https://www.mathworks.com/products/matlab.html
  For licensed version in LUMC request via the IT department

- Ilastik (any version)
  Used for training of probability masks
  Can be downloaded from www.ilastik.org/download

- Cellprofiler (version 2.2.0)
  Used to convert probability masks from ilastik to masks
  Can be downloaded from https://cellprofiler.org/previous_releases/
  Software needs java 64-bit to install which is only installed if the browser used to download is 64-bit.

- Imacyte
  Used for the analysis of images, phenotype identification, neighbourhood analyses and cohort comparison
  Can be downloaded from https://github.com/biovault/ImaCytE

- Cytosplore (optional)
  Used for clustering and phenotype identification of large datasets
  Can be downloaded from www.cytosplore.org

Overview of the pipeline (Figure 1)

1. Conversion of .mcd files to single marker .ome-tiff files in MCD viewer

   Raw datafiles from the Hyperion are saved as the .mcd extension which is not readable by other programs. Therefore the files have to be converted to .ome-tiff files.

2. Image enhancement in Matlab

   To facilitate pixel annotation, the images are sharpened in Matlab. More specifically, outliers are removed through saturation of all pixels with values lower than the 1st and higher than the 99th percentile. Furthermore, as most software can not read .ome-tiff files, the files are changed to a standard tiff format.

3. Background-removal for each marker by Ilastik

   To remove background and variation between samples the images are normalized. This is done by loading all images of a single marker into Ilastik and training a random forest classifier to recognize what is background and real signal. This is applied to all images of a certain marker in the dataset and exported as binary masks. All background pixels are set to 0 and all real pixels are set to 1.

4. Training of masks in ilastik and creation of masks in cell profiler

   To obtain single cell data the software must know which pixels belong to which cell. To do this a cell mask to overlay the images is created using Ilastik and cell profiler. A training image is made in mcd viewer after which ilastik is trained to recognize which pixels of the image belong to the background, membrane/cytoplasm or nuclei. This information is exported as a probability mask. To create the final mask, cell profiler uses the probability mask as input and combines the pixels assigned as membrane/cytoplasm and nuclei to create the cells. Furthermore it filters out single pixels and splits cells that are too big.

   The steps described in this manual work on most tissues, however for very dense tissues an adapted approach might be needed. For instance, performing cell segmentation in 2 steps where first the myeloid cells are segmented and afterwards the T-cells. In this case a different CellProfiler pipeline can be used than the one described here.

5. Data visualization and inspection in ImaCytE

   The data can be loaded into ImaCytE, which allows the exploration of marker expression in the exploration of marker expression in some or all images in the dataset.

   Furthermore, ImaCytE can be used for direct TSNE clustering of the data or single cell expression data can be exported as FCS or CSV files for analysis with external software. After phenotype identification ImaCytE allows the analysis of spatial localisation and interactions of phenotypes in the whole cohort or via cohort comparison.

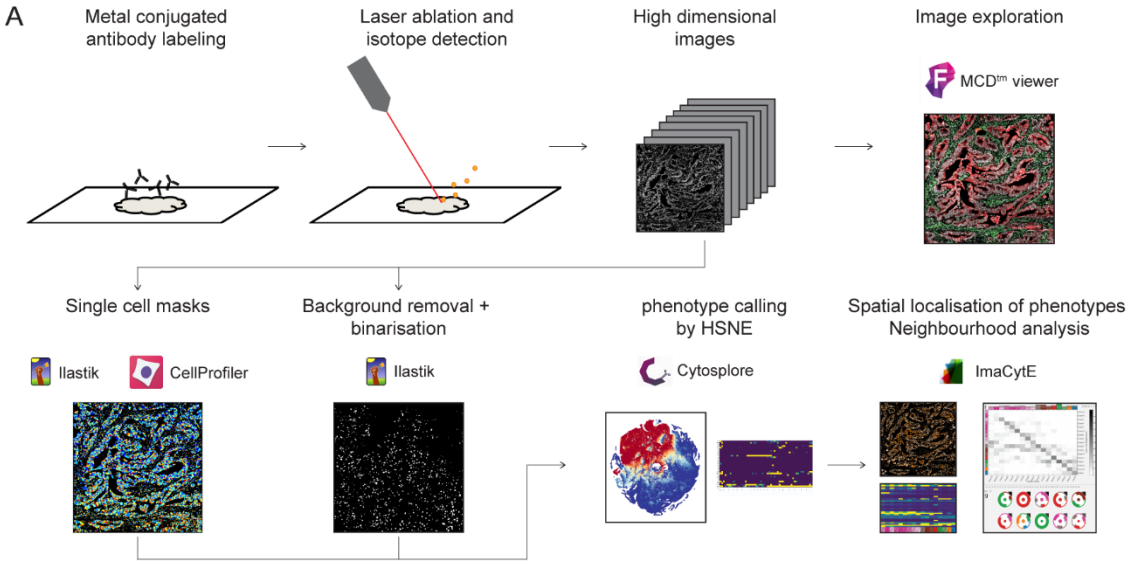

Figure 1| Workflow for Imaging mass cytometry analysis

Step by step IMC analysis

It is recommended to read the overview of the pipeline before starting. In the following guide we will explain how to convert data for IMC analysis, perform background removal in ilastik and analyse data in ImaCytE. The creation of masks is not addressed as these can be created in various ways depending on the tissue and markers.

IMC analysis using raw data

Conversion of .mcd files to .ome-tiff files in MCD viewer

1. Open the .mcd file and load one of the ROIs
2. Go to 'File' --> 'export'
3. Change export type to OME-TIFF 32-bit and page to multi
4. Choose an export location in the 'file name' box (Figure 2)
5. Select the ROIs to export and click 'export'

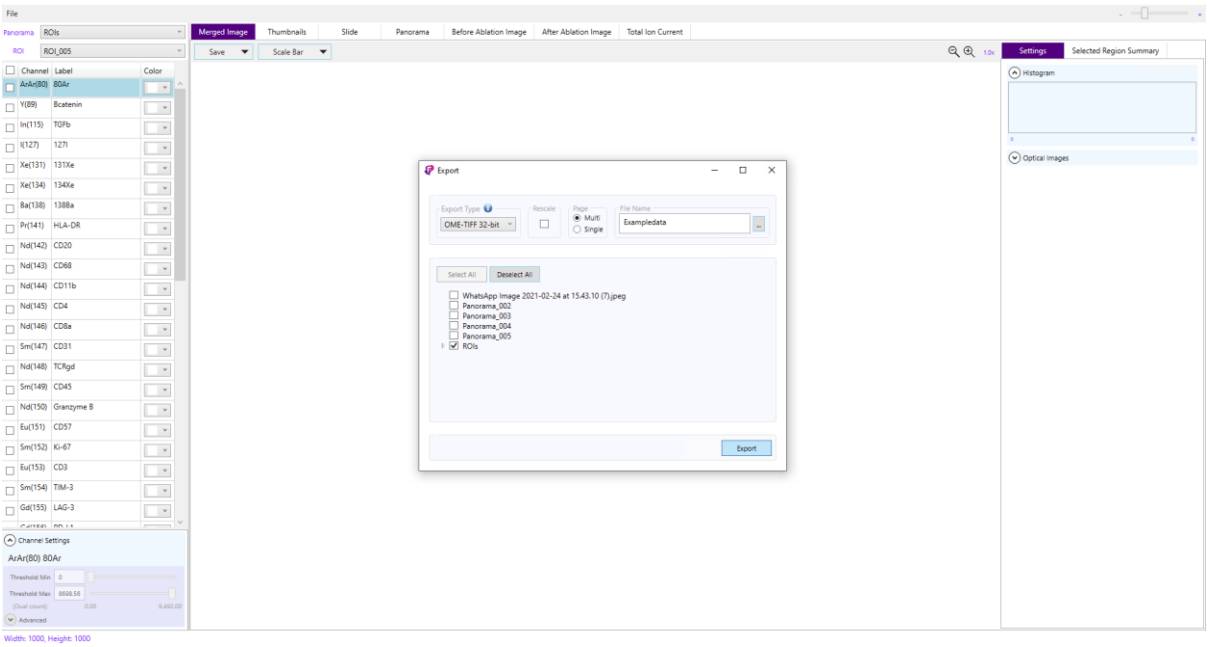

Figure 2|Settings for export of .ome-tiff files

Cell segmentation mask creation

As masks can be created in various ways which differ between tissuetypes and cells of interest, this is not covered in the current guide. Masks from various methods will work with ImaCytE.

Organisation of data to load into ImaCyte

6. For ImaCyte to load the data properly a folder is created for each ROI. If the manual is followed this was already done when exporting the ome.tiff files (step 5). Confirm in the next steps if all files are correct or create the folders manually.
7. Create a folder and name it "ROIs'
8. Place all folder containing the ROIs inside this folder (Figure 3)
9. Each ROI folder should contain (Figure 3):

For unnormalized data analysis

a. Ome.tiff file
b. Mask file (uint16, name must end in _mask_1)
c. Tiles file

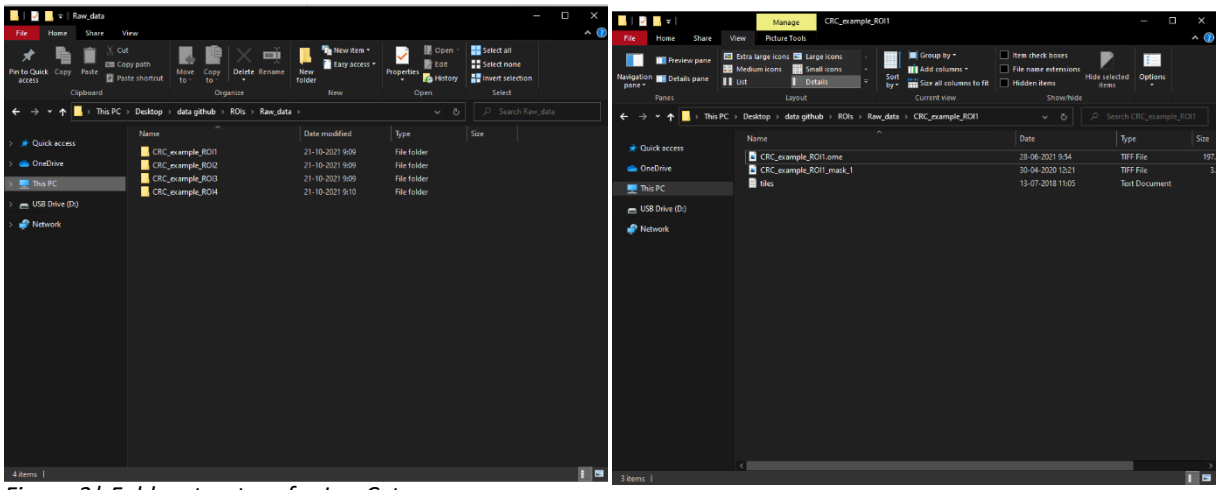

Figure 3| Folder structure for ImaCyte

10. Continue with step 31

IMC analysis on normalised data using ilastik background removal

Conversion of .mcd files to .ome-tiff files in MCD viewer

11. Open the .mcd file and load one of the ROIs
12. Go to 'File' --> 'export'
13. Change export type to OME-TIFF 32-bit and page to single
14. Choose an export location in the 'file name' box (Figure 4)
15. Select the ROIs to export and click 'export'

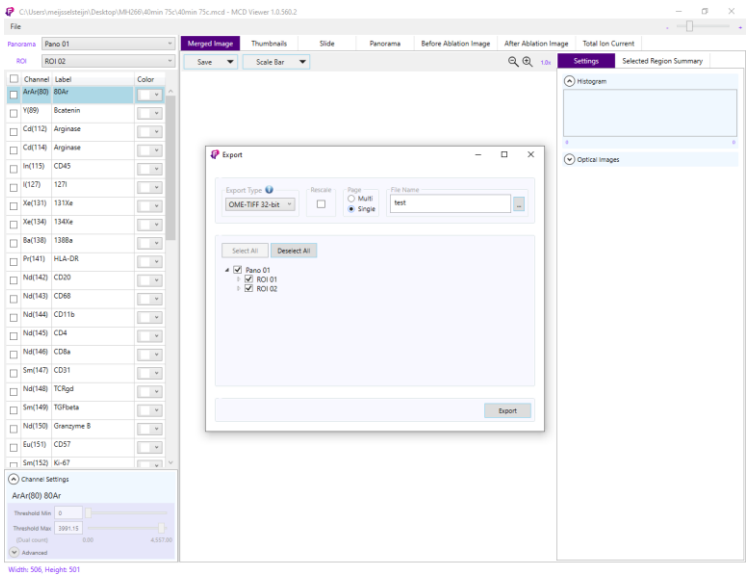

Figure 4|Settings for export of .ome-tiff files

(optional) File renaming

To convert the files using the matlab script in the next section the images are converted per ROI. By adding the ROI name to the marker images, all images can be converted at once.

File conversion to load into Ilastik for background-removal

16. Open the Prepare_Images_for_Ilastik3.m script in matlab
17. The script makes use of two folders: the 'ome' folder and the 'output' folder. Use the images exported in step 15 and place them in the 'ome' folder
18. Assign the input and output folders and run the script. the enhanced images will be created in the output folder

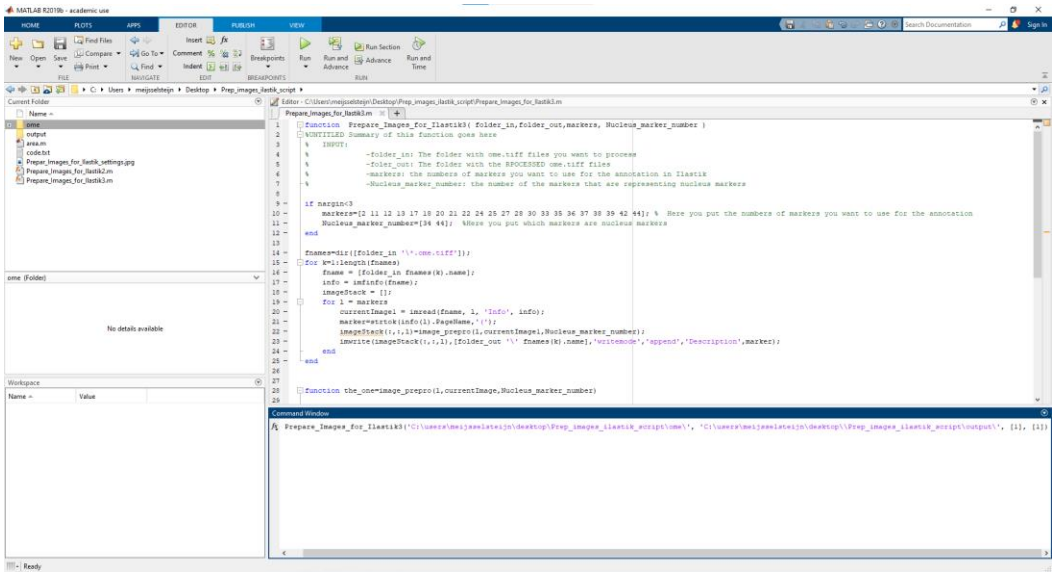

Figure 5| Matlab script to transform datatype and enhance images

Background-removal for each marker by Ilastik

19. Open Ilastik project Ilastik_Backgroundremoval.ilp
20. For the entire dataset, load all images corresponding to one marker by dragging the images into the white box (Figure 6)
21. Continue to '3. Training' and draw in pixels corresponding to ' background' and 'signal' (Figure 7)
22. Tick ' Live update' and 'segmentation' to check if pixels are assigned properly (Figure 8).

    Note: on a fast computer 'Live update' can stay on, otherwise tick off when drawing'
23. Assign pixels until real signal and background are properly segmented in multiple images.
24. Continue to '4. Prediction export' and confirm that 'source' is set to simple segmentation.
25. Export all Threshold masks by clicking 'Export all'

    Note: it is essential that the filenames end in _threshold.tiff or ImaCytE will not recognize the normalisation
26. Repeat step 19-25 for all markers of interest

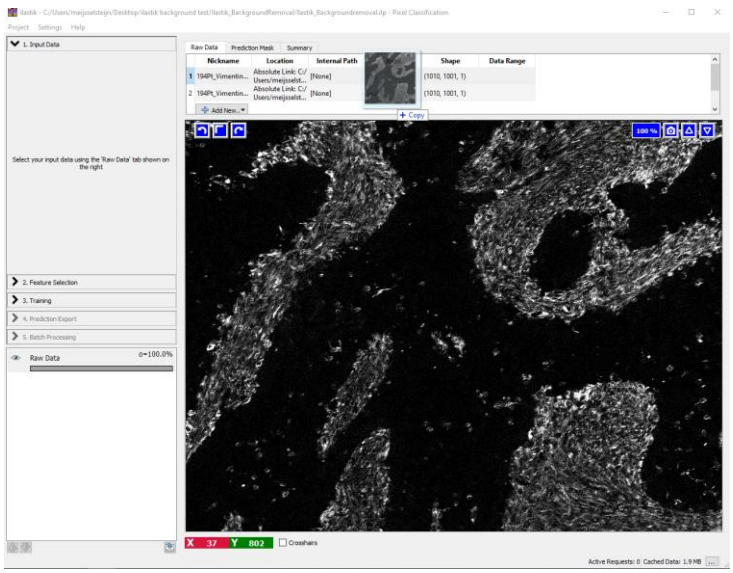

Figure 6| Loading images into Ilastik

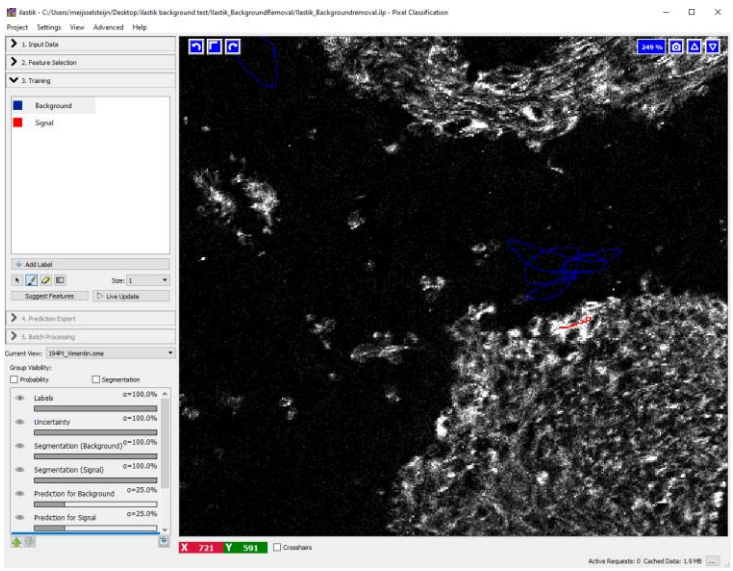

Figure 7| Ilastik: Assigning pixels to background or signal for one marker

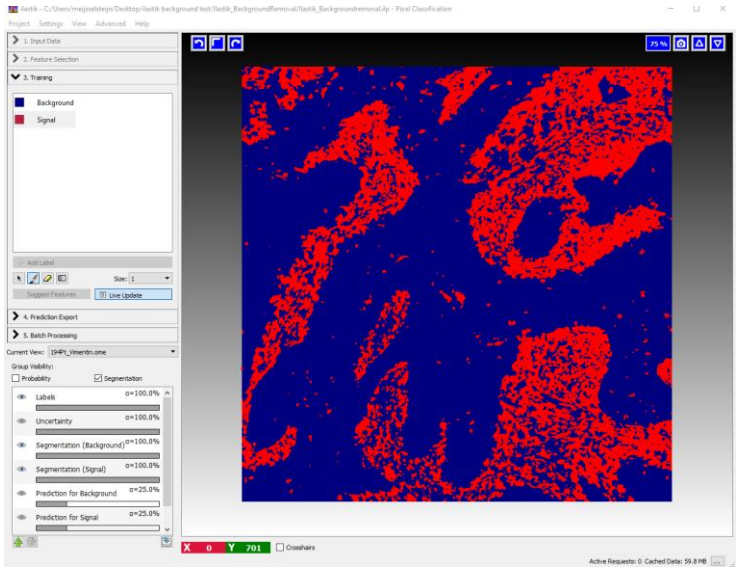

Figure 8| Ilastik: live update of pixels assigned to background or signal

Organisation of data to load into ImaCyte

27. For ImaCyte to load the data properly a folder is created for each ROI. If the manual is followed this was already done in step 5 which exported the ome.tiff files. Confirm in the next steps if all files are correct or create the folders manually.
28. Create a folder and name it “ROIs’
29. Place all folder containing the ROIs inside this folder (Figure 9)
30. Each ROI folder should contain (Figure 9):

    For normalized data analysis

    a. Mask file (uint16, name must end in _mask_1)
    b. Threshold files for each individual marker of interest (tiffs, names must end in _threshold)

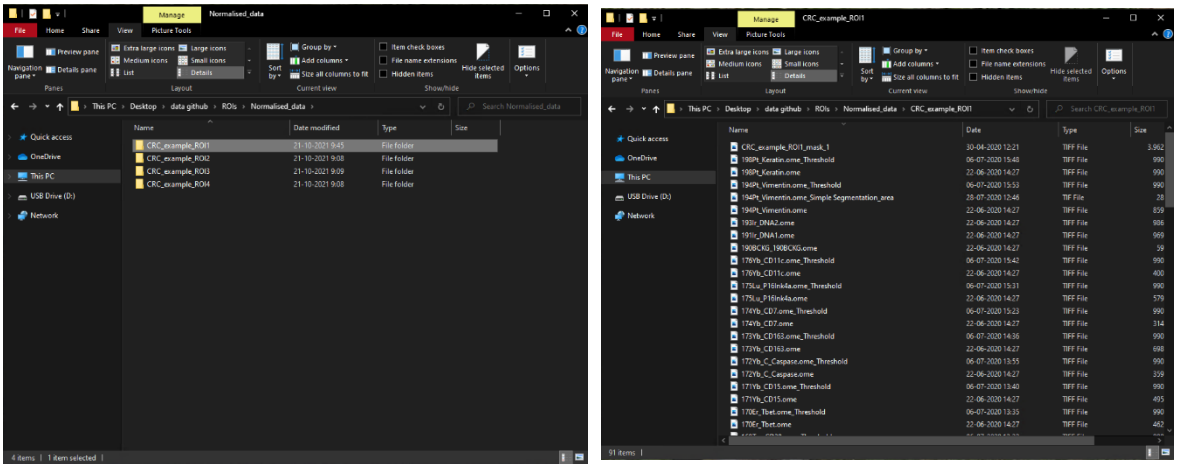

Figure 9| Folder structure for ImaCyte

Data visualization and inspection in ImaCytE (figure 10-18)

31. Run ImaCytE as administrator
32. Click 'Options' → 'Load data'
33. Choose if you want to load raw images, thresholded values or normalised data
34. Open the 'ROIs' folder and click 'Select folder'
35. ImaCytE will show a loading bar and load the images
36. Visually inspect the images by selecting 'samples' and 'markers'.

    Note: you can either choose several samples or markers, not both

37. You can now either run a TSNE analysis directly or export the single cell information via 'Options'→ 'export data per sample' → 'as fcs' or 'as .csv' and perform phenotype identification with the method of choice

38. Phenotype FCS files can be reloaded into ImaCytE via 'Options' → 'Load phenotypes' → 'per phenotype (.FCS)'

39. After reloading, FCS files can be merged further if needed and the session containing these can be saved, which can be reloaded.

40. Interaction analysis can be performed on the whole dataset or selected images via the 'interaction analysis' button.

41. Two groups/cohorts can be compared using the 'cohort comparison', which allows the exploration of 'cell abundances' and 'cell colocalization' via the drop down in the pop up window

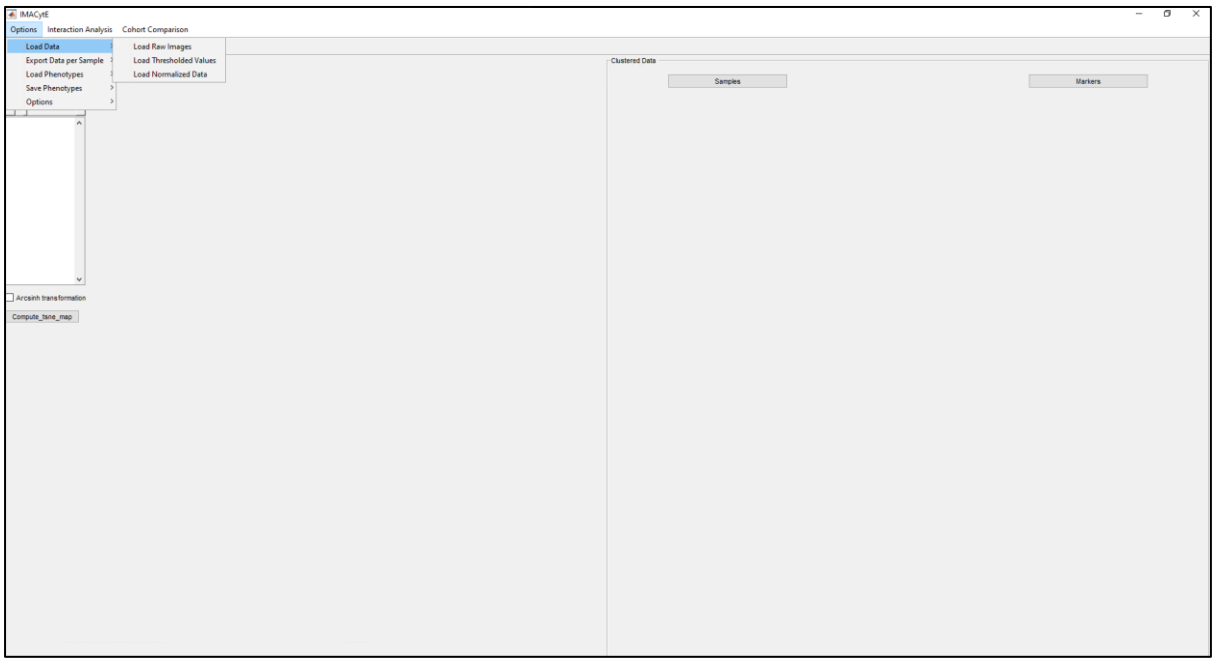

Figure 10|ImaCytE: choose which data to load

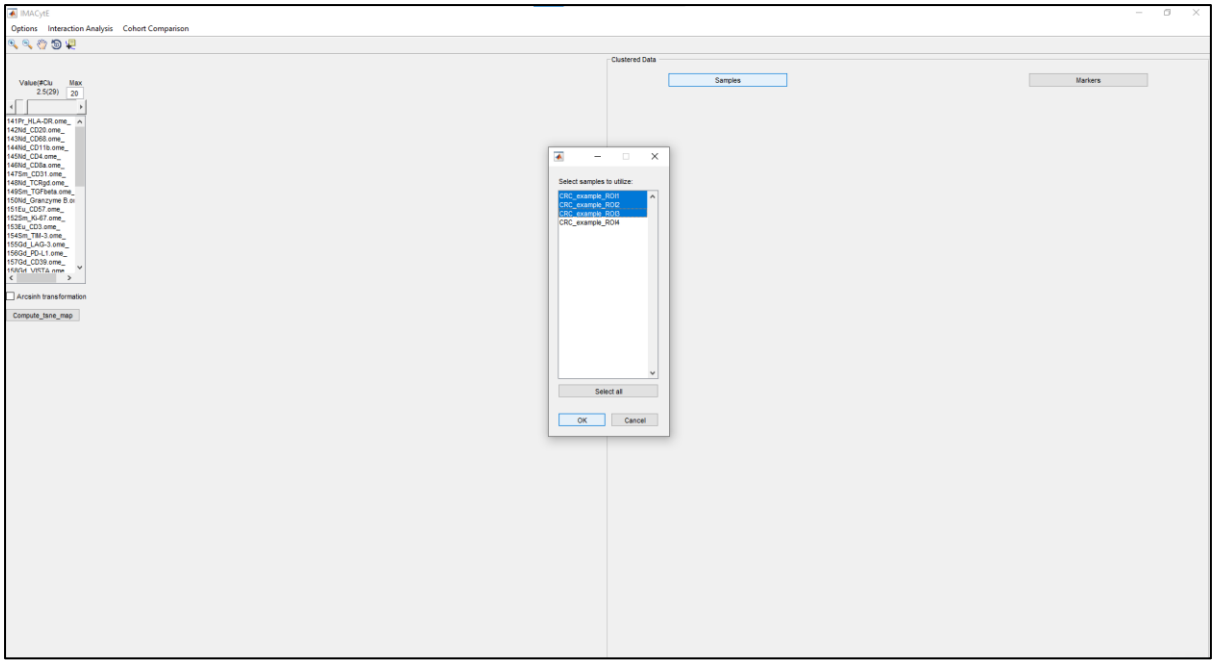

Figure 11|ImaCytE: select images to visualize

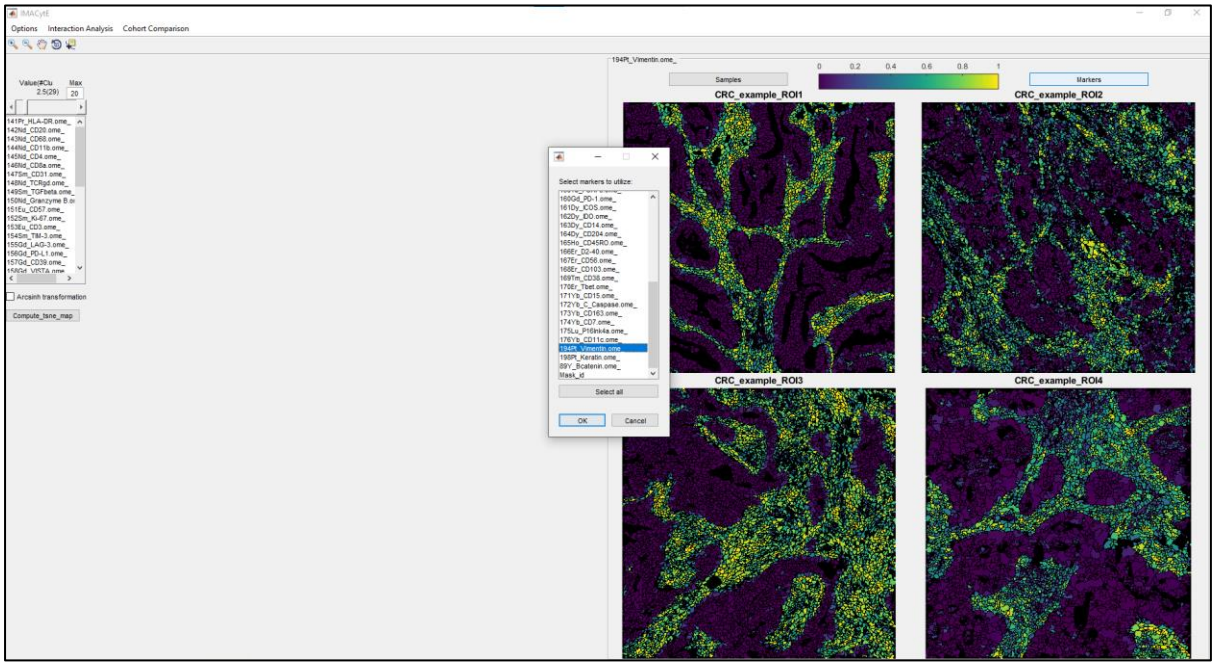

Figure 12|ImaCytE: select markers to visualise

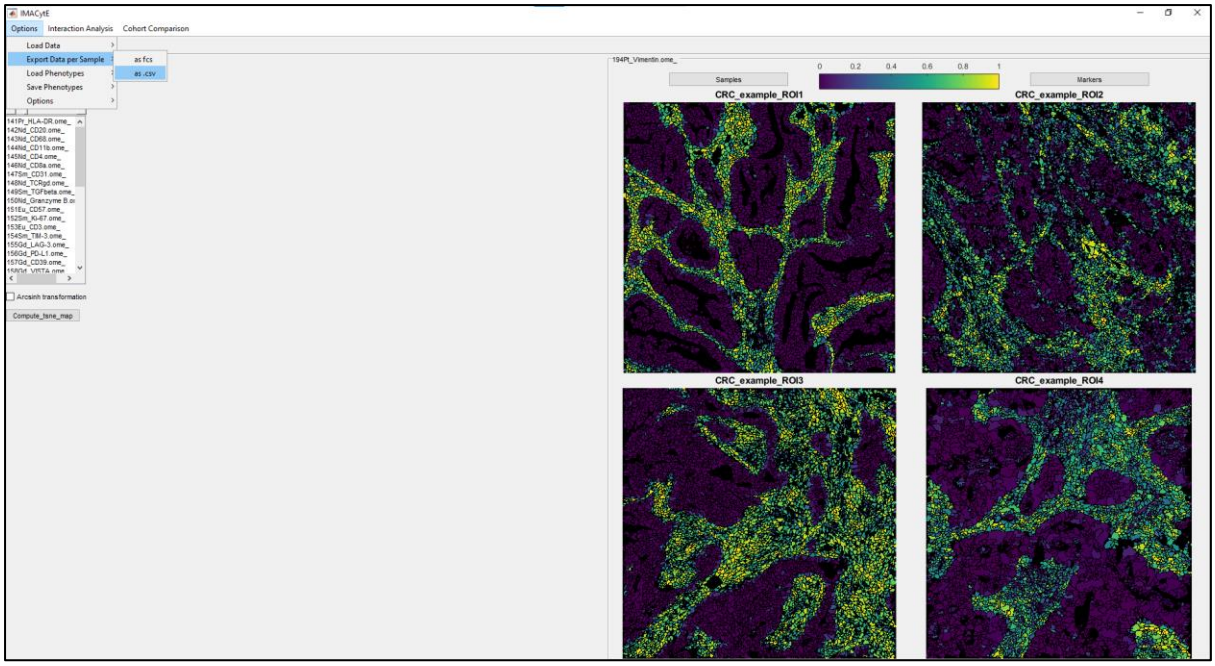

Figure 13|exporting single cell information as CSV or FCS

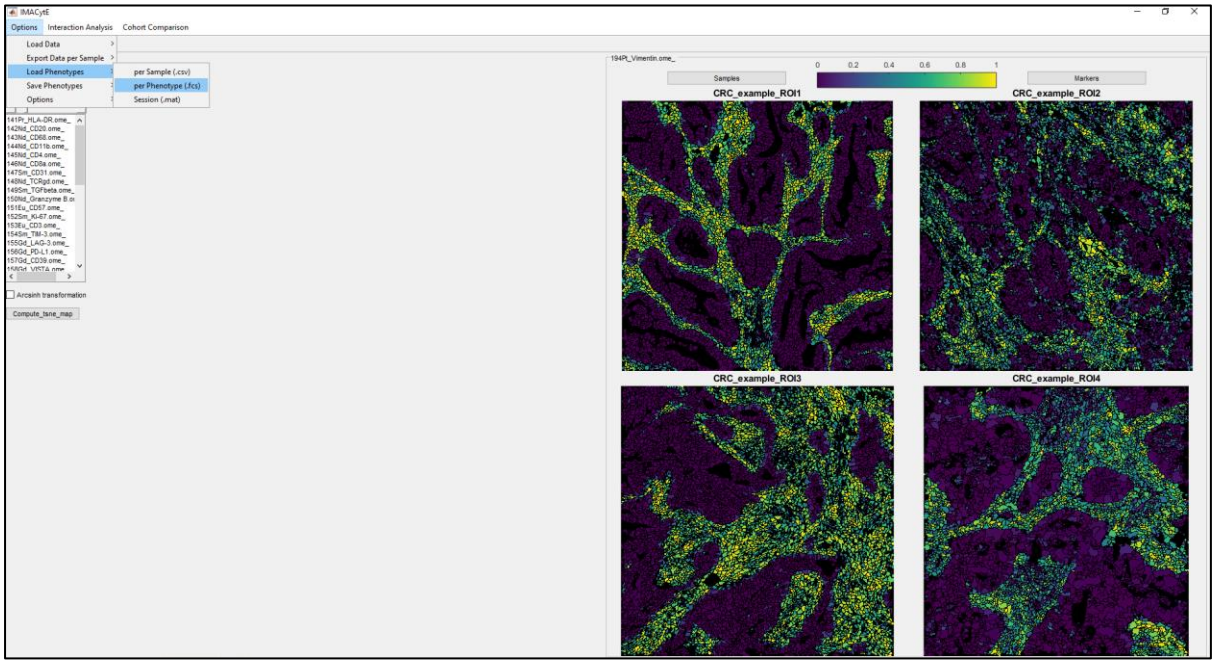

Figure 14|importing phenotype files (.fcs)

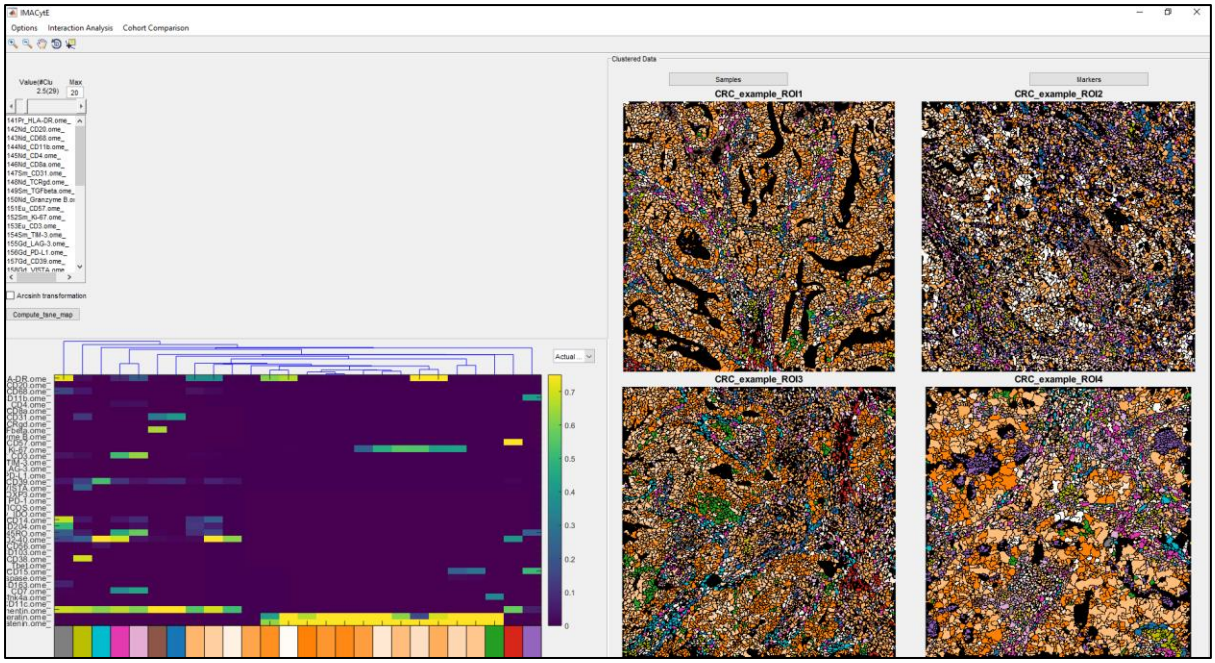

Figure 15|the phenotypes will be put into a heatmap and mapped onto the images

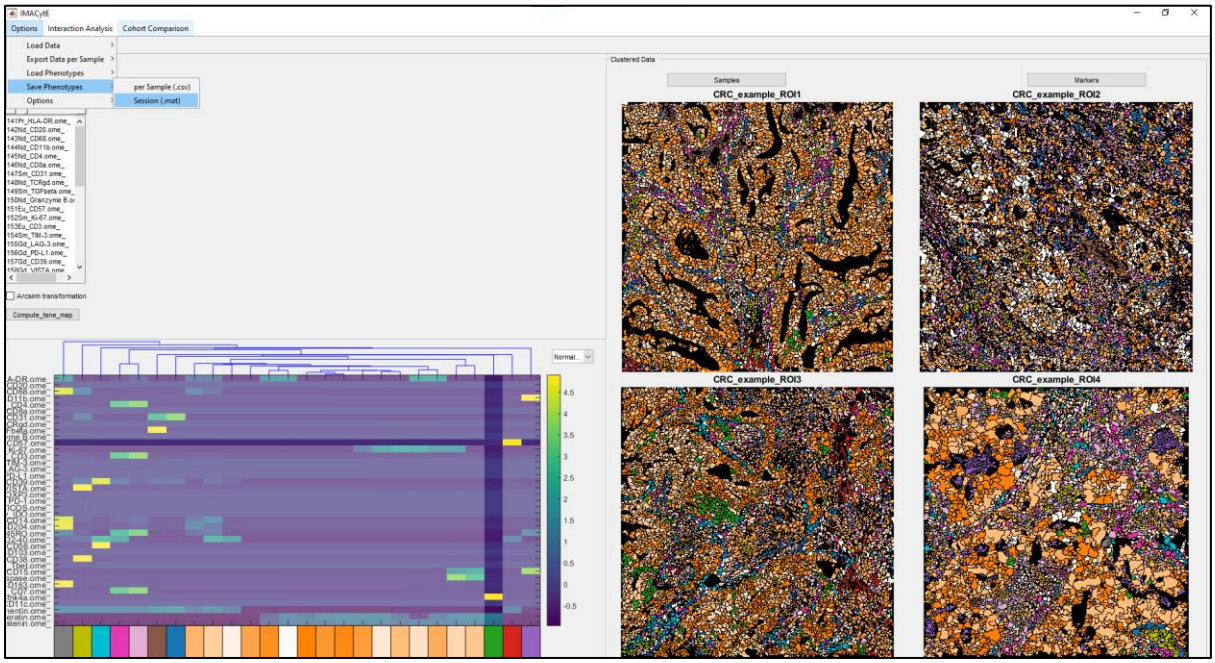

Figure 16|saving the heatmap/phenotype session

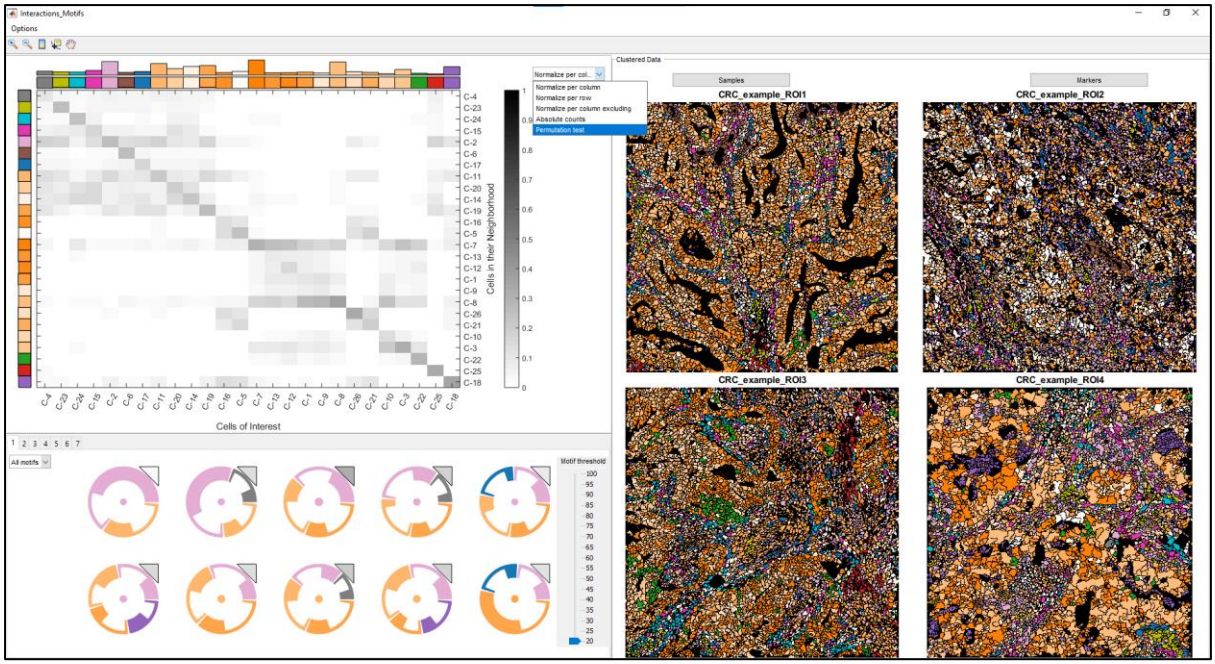

Figure 17|Interaction analysis

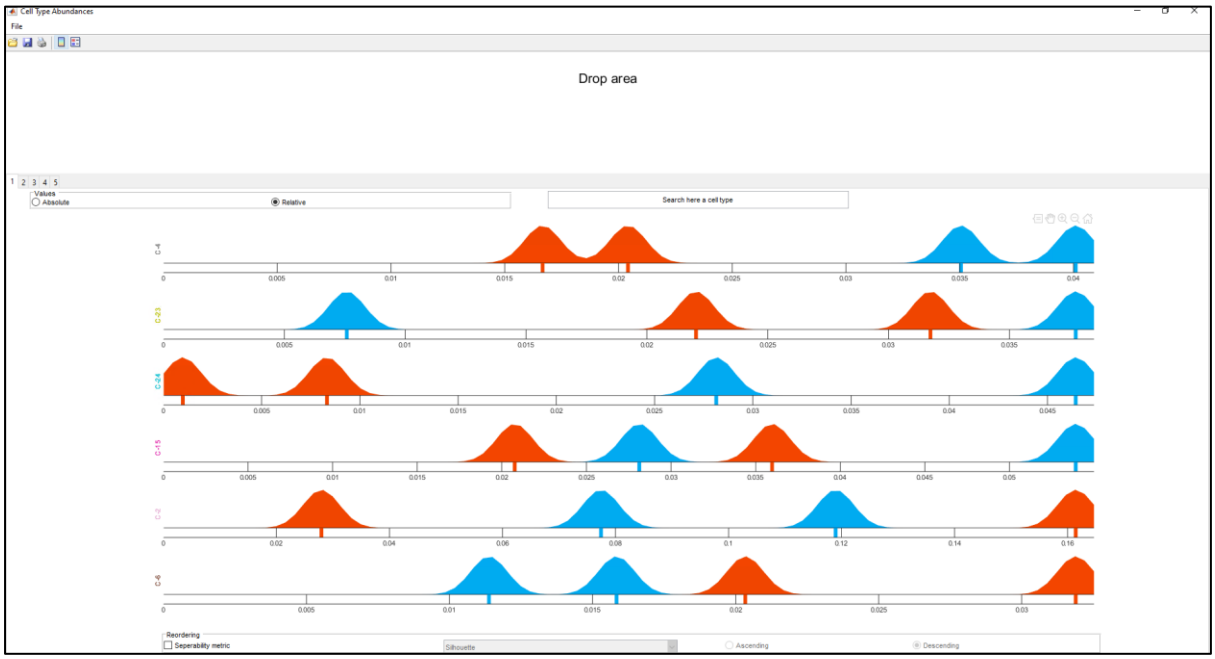

Figure 18|cell type abundances in cohort comparison

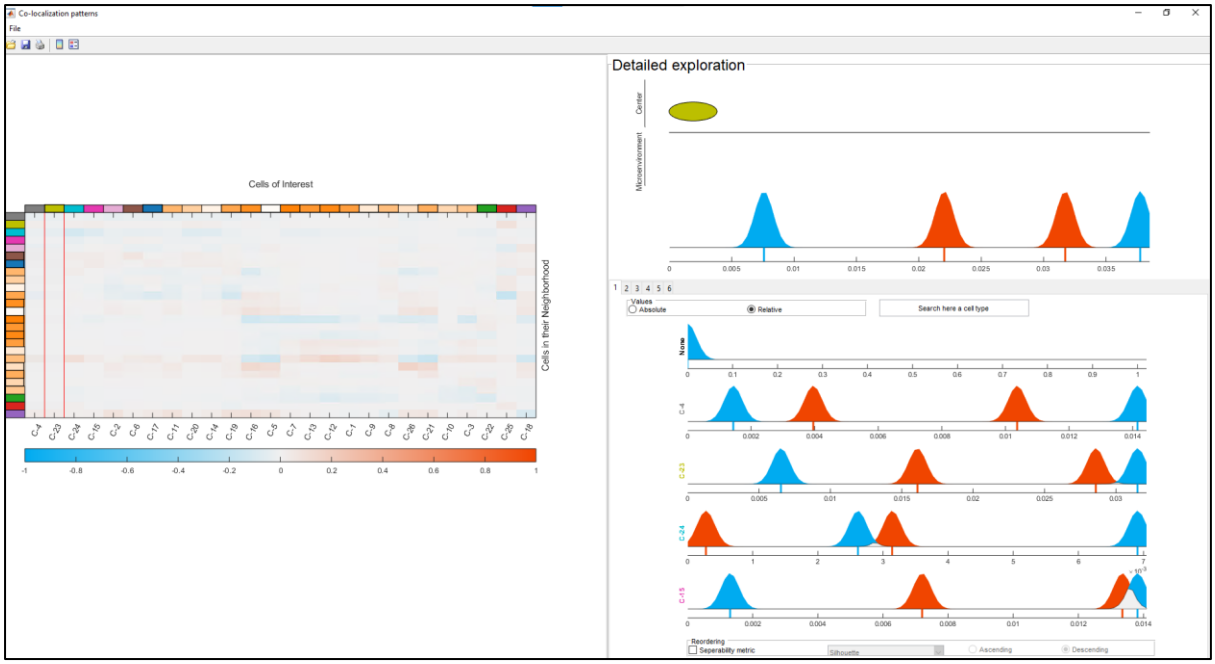

Figure 19|colocalisation comparison in cohort comparison