Single-Cell spatial Explorer Manual

Frédéric PONT

February 2, 2022

Contents

0.1 Single-Cell spatial Explorer presentation

Single-Cell Spatial Explorer is a new tool, part of Single-Cell Signature Explorer software package [1] dedicated to spatial transcriptomics. It is able to display single cell data using color gradients on a microscopy image. It is possible to

define and gate cell populations based on the cells or the single cell data. An unlimited number of data tables containing XY image coordinates can be filtered according to the gates. The content of the gates can be plotted with any XY coordinates or compared. Gates comparison is done by plotting an interactive vulcano plot, in which it is possible to select points of interest and display the corresponding gene/pathway expression.

Files can be accessed at https://sites.google.com/site/fredsoftwares/products/single-cell-signature-explorer

0.1.1 Computer configuration.

Minimal computer configuration.

Single-Cell Spatial Explorer has been developed on linux Manjaro on a computer with an Intel ©Core™i5-3470 CPU @ 3.20GHz with 8 Go RAM, a 7200 rpm hard disk, a video card Nvidia GeForce GT630 and a HD display with a resolution 1920×1080. This is a comfortable configuration for a small data set of about 3000 cells.
Although it could be possible to use Single-Cell Spatial Explorer on a laptop, we recommend to use a HD display.

Recommended computer configuration.

Hard disk

Single-Cell Spatial Explorer has been optimized to save RAM at the expense of hard disk access. As a result, Single-Cell Spatial Explorer will take a great advantage of fast disks such as SSD disks. Copy the software folder on your fastest disk or RAID.

CPU

Except tables filtering which can be massively parallel (1 table / CPU), most of the tasks uses 2-3 CPU. Single-Cell Spatial Explorer will be fast on CPU with a high single thread speed, such as the CPUs on top of this list.

Memory

The gates comparison is the most RAM demanding operation because the whole dataset have to fit in RAM if the user select all the columns in the comparison. For 3000 cells and 20 000 genes, 8Go RAM are enough (at least on a light Linux distribution), but for larger data sets, it could be necessary to increase RAM over 8Go.

Display

The display is very important for an optimal use of Single-Cell Spatial Explorer. We recommend at least one HD display, two if possible, or a high resolution

display such as 4k displays. With a 4k display, it is possible to work on a 2000×2000 image at zoom 100% which is the more precise zoom since there is no pixel interpolation. Very important : do not use display zooming, use the screen at the native resolution with a 100% zoom.

0.1.2 Single-Cell Spatial Explorer features.

1. Single-Cell Spatial Explorer is ready to use in a pre-compiled binary, no installation required
2. cross-platform (the interface and the software are coded in pure Go)
3. low memory usage
4. compatible with any PNG image associated with any CSV file (TAB separated) containing XY coordinates of the image.
5. compatible with any numeric data : gene expression, pathway scores, antibody expression etc...
6. unlimited number of gates.
7. import/export gates in ImageJ/Fidji format.
8. extract cells and sub-tables delimited by the gates on an unlimited number of tables. Exportation is done in CSV file for great interoperability.
9. 2D plots of the cells inside the gates with any XY coordinates : t-SNE, UMAP, gene expression, pathway scores, antibody expression etc...
10. interactive 2D plot to show the selected cells on a t-SNE, UMAP or any other coordinates on the image and to filter the data tables into sub-tables.
11. cluster display with custom dot opacity and custom dot size.
12. display any kind of cell expression (genes, pathways, antibodies...) with 7 preset gradients, custom legend color, dot opacity and custom dot size. The gradients are simple two colors maps and rainbow colors maps Turbo, Viridis and Inferno to optimize accuracy and details visualisation.
13. Min/Max intensity sliders to tune image contrast or remove artefacts due to outliers.
14. slide show to review many cell expression maps without need of repetitive click.
15. screenshot or native resolution image exportation.
16. import and display an unlimited number of cells list by repetitive click on the "import cells" button. The format is directly compatible with Single-Cell Virtual Cytometer [2].

17. compare two groups of gates together in the whole dataset .
18. compare one group of gates against all the remaining cells.
19. draw an interactive vulcano plot after gate comparison.
20. plot cell expression of a selected dot in the vulcano plot.
21. export vulcano plot and the corresponding data table.
22. image zoom 10-200%

0.2 Software installation and use

Single-Cell Spatial Explorer has been compiled for GNU Linux and Microsoft©Windows™64 bits and require no complex installation procedure. A standalone executable file requiring no dependencies is provided.

0.2.1 Quick start

1. unzip the software
2. we recommend to test the software with the test files (data table + image) first and to save a copy of the test files.
3. copy Single-Cell data csv/tsv file to the "data" directory : the first row are gene or pathway (or anything else) names, the first column are cell barcodes and the next columns are the expression values, with separator TAB. The table MUST contain [X,Y] coordinates of the cells in the microscopy image. This is an example of the table :

id	P53_PATHWAY	X_hires_image	Y_hires_image
cell_AAACTGC	2.083	13378	4022
cell_AAGACCT	2.211	4340	12932
cell_AACAGAG	2.397	10102	2899

It is possible to copy an unlimited number of data tables in the "data" directory : the first table by alphabetical order with be used by the software, but all the tables can be filtered by gates.

4. copy the microscopy image in PNG format in the "image" directory
5. start the software using the binary file matching your operating system. Do not start multiple instances of the software. Multiple instance will try to use the same temporary files, such as the parameters preferences. This could give unexpected results or worse corrupt the preference file. After using any tool box, except the main panel of the software, is it highly recommended to close it, before opening another one.

6. click on the "Preferences" button and set all the preferences. Close the "Preferences" window and open it again to verify that all the settings are well saved.

7. once the preferences are set, restart the software.

8. Very important : when you display clusters/expression, do not forget to set the opacity slider between 50-100% opacity or you will see no results for clusters or expression.

0.2.2 Preferences setting

It is mandatory to set the preferences first before working with the software and then restart the software. Indeed the link between the microscopy image and the data table is made using column names and values that are in the preferences. Preferences items :

- Scaling Factor : the scaling factor is the multiplicative number making the link between the [X,Y] coordinates in the data table and the image.
- Rotate : if the coordinates in the table are rotated ($-90°$) compare to the coordinates of the image, check this option
- X,Y coordinates : the columns of the data table with the [X,Y] coordinates of the image
- Image window width/height : the size of the image window. CAUTION the window size must fit the display. For example for a HD screen, the maximum resolution is 1920×1080 and the maximum recommended window size is 1000×1000.
- Cluster column : the column of the data table with the cluster number. This column must contains only integer numbers. Each integer is used as a cluster number.
- Cluster dots diameter : the point size in pixels to display clusters or expression.
- Vulcano selection square size in pixels : the vulcano plot selection square size in pixels. This square is used to select dots in the vulcano plot

0.2.3 Drawing gates on the microscopy image

Single-Cell Spatial Explorer features can work with an unlimited number of gates.

- Select the brush or the polygon. The selected tool is green.

- To draw a gate with the brush, left click on the image with the mouse and drag the mouse. Then release the mouse button. The gate is automatically closed between the beginning and the end, it is not necessary that he beginning and the end are at the same place.
- To draw a gate with the polygon, click on the image, a dot appears, move the mouse and click again. To close the shape, right click.The gate is automatically closed between the beginning and the end, it is not necessary that he beginning and the end are at the same place.
- To clear all gates, click on the "clear all gates" button
- To clear the last gate, click on the "clear last gate" button

0.2.4 Exporting gates

Gates are exported in CSV format in the "gates" directory using the "Save gates" button. The name of the gate is the number of the gate, followed by the name in the text entry of the tool window. Gates are also automatically exported when the tables are filtered by gates. cf. 0.2.5

Importing gates

Single-Cell Spatial Explorer can import an unlimited number of gates. A gate file is a CSV file with a X,Y header followed by the edges coordinates of the polygon.
X,Y
851,891
1149,833
1171,1073
This format is compatible with ImageJ/Fidji image analysis software.
The gates files have to be copied in the "import_gates" directory and then imported using the "Import Gates" button

0.2.5 Filtering data with gates

Single-Cell Spatial Explorer can filter an unlimited number of data tables in the "data" directory using the "Filter tables with gates" button. The filtered tables and the cell names are in the "results" directory and the gates are automatically saved in the "gates" directory using the name in the text entry of the tool window. The names of the sub-tables can be set in the text entry on the top of the 2D plot tool window. Two types of files are produced : "cells" files contain the cell names and "filtered" files contain the sub-tables. The file names is "cells/filtered" _ gate number _ text entry _ orginal table name. In the sub-tables, a column labeled "GateNumber" is added at the end of the table.

0.2.6 Plotting gates

The cells in gate can be plotted with any 2D coordinates using the "Plot gates" button. Do not forget to set the dots colors and the dots in background colors. If the dots colors is white, they will not appear in the scatter plot. An optimized set of colors is used first. If the number of gates exceed the number of preset colors (10), then the "Turbo" gradient is used to compute an unlimited number of colors through the rainbow. The plot are saved in the "plots" directory.

0.2.7 Comparing gates

Gates and groups of gates can be compared using the "Compare gates" button. This button opens a "Compare" window. The columns used for the comparison can be selected using check boxes. Cells inside two gates or two groups of gates can be compared using the "Compare Group2 / Group1" button. Cells inside a gate or a group of gates can be compared with the cells outside the gates using the "Compare Outside / Group1" button. These buttons open an interactive vulcano plot and a vulcano tool window. To save the vulcano plot, click on the "Save vulcano plot" button. The vulcano plot data are automatically saved in the "comparison" directory. A left click on the vulcano plot display a selection square (the size is set in preferences). A right click list the dots names, fold change and pvalue in the vulcano tool window. In the vulcano tool window, it is possible to select a color gradient, a dot opacity (do not forget to set the dot opacity between 50-100% to see the dots) and a dot in the item table. The "Show expression" button will display the expression associated with the vulcano dot on the image.

0.2.8 2D interactive plots

It is possible to display an interactive scatter plot by clicking on the "2D interactive plots" button. Any 2D coordinates can be used, such as t-SNE, UMAP, gene expression, antibody signal etc... It is possible to draw gates on the 2D interactive plots. Left click on the image with the mouse and drag the mouse. Then release the mouse button. The gate is automatically closed between the beginning and the end, it is not necessary that he beginning and the end are at the same place.

1. Once the gates are drawn, it is possible to display the selected cells on the image by clicking on the "Show Cells in Gates" button.

2. Using the button "Filter tables by Gates", an unlimited number of data tables, in the "data" directory can be filtered to extract the sub-tables corresponding to the cells in the gates. The gates are automatically saved in the "gates" directory. The names of the sub-tables can be set in the text entry on the top of the 2D plot tool window. see 0.2.5.

3. To save gates (without filtering the tables), use the "Save Gates" button.

4. To import gates, use the "Import Gates" button. The gates must be copied in the "import_gates_2Dplot" directory. A gate file is a CSV file with a X,Y header followed by the edges coordinates of the polygon.
   X,Y
   851,891
   1149,833
   1171,1073
   This format is compatible with ImageJ/Fidji image analysis software.

5. To clear gates, use the "Clear Gates" button.

6. To save a screenshot of the 2D plot press the camera button.

0.2.9 Displaying cluster

To be able to show clusters on the microscopy image, it is necessary to identify the clusters column in the preference setting. Then press the "Show Clusters" button. The colors gradient can be selected in the preference setting. When "Shuffle colors" is check, a new distribution of colors is obtained following repeated click on the "Show Clusters" button.
The color palettes are saved in the "palettes" directory, thus it is possible to use the colors in another plotting software. For very experienced users, an unlimited number of palettes can be placed in JSON format in src/palette. We do not recommend to modify the src/palette/palette1.json file. It is better to make a copy of this file and work on the copy. A software restart is required to make custom palettes appear in the preference form.
With 12 clusters, almost 2 billions of different images can be obtained with the 4 gradients and the shuffle colors.

0.2.10 Displaying expression

Expression can be genes expression, antibody signal, gene sets expression, or any numerical data present in the data table.
To display expression data on the microscopy image, press the "Expression" button. A window open :

1. on the top of the window, select a column of the data table in the drop down list and press "submit".

2. select a color gradient

3. if necessary change the legend color using the "Legend Text Color" button

4. press the "Plot Expression" button

5. the expression is displayed on the microscopy image. The dot opacity, and the image contrast can be adjusted with the sliders "Dots opacity", "Min", "Max". Press the "Apply" button to apply changes.

6. when the check box "Opacity gradient" is selected, the opacity for each spot is computed from the expression value. The min expression value correspond to a full transparency and the max a full opacity. The min/-max opacity threshold must be entered in the entry spaces. For example, if min opacity = 1.5, the spots with an expression ¡ 1.5 are fully transparent. if max opacity = 3.5, the spots with an expression ¿ 3.5 are fully opaque. When min opacity is closed to max opacity a kind of binarisation is obtained and only the spots with an expression between min and max are visible. min opacity must be superior to max opacity otherwise the software ignore the opacity computation.

7. In the expression window, a 2D plot show the distribution profile of the dots intensities. This plot is obtained by dividing the expression levels into 100 classes that are smoothed by a moving average on 20 points.

Slide Show

To do a slide show :

1. Enter the pause between slides in the "Pause between slides" entry

2. Press the "Slide show" button

3. to stop/continue the slide show : press the "stop/continue slide show"

4. to show next/previous slide : press the "Next slide" / "Previous slide" buttons

0.2.11 Importing Cells

An unlimited number of files containing cell names can be placed in the "import_cells". The files must be text files (csv, txt) with one cell name per line without header. import and display an unlimited number of cells list by repetitive click on the "import cells" button. The format is directly compatible with Single-Cell Virtual Cytometer

0.2.12 Troubleshooting

The software crashes when I display clusters or expression.

It is very important to test first the software with tests files and the preferences settings with the test files. If the software works with the test files and crashes with other data files, the possible causes are :

1. preferences are not set for the data file, some columns cannot be found

2. the data table is not formated as required or contains special characters, blanks or strings instead of numbers

There is a shift between the image and the cluster/expression map.

1. the scaling factor is wrong
2. the coordinates in the data file do not match the pixel of the image. Check the resolution of the image and the coordinates column in the data file.

My scatter plot is not displayed or partially displayed

1. Check the colors of the dots in gate and dots in background. A white spot is not visible over a white background.

Bibliography

[1] Pont, F., Tosolini, M., and Fournié, J. J. (2019) Single-Cell Signature Explorer for comprehensive visualization of single cell signatures across scRNA-seq datasets. *Nucleic acids research,* **47**(21), e133–e133.

[2] Pont, F., Tosolini, M., Gao, Q., Perrier, M., Madrid-Mencía, M., Huang, T. S., Neuvial, P., Ayyoub, M., Nazor, K., and Fournié, J.-J. (2020) Single-Cell Virtual Cytometer allows user-friendly and versatile analysis and visualization of multimodal single cell RNAseq datasets. *NAR genomics and bioinformatics,* **2**(2), lqaa025.