Multiparametric Flow Cytometry First Steps

Carine P. Beatrici

January 24, 2018

1 Introduction

The multiparametric-flow-cytometry is the program used to generate the figures in the paper "A high-throughput computational tool to analyze phenotypic modulations in multiparametric flow cytometry data". Here we present how to use the program to reproduce the same results as well to get new results with any flow cytometry standard file (FCS).

To compile and execute the program the dependencies are:

1. A Linux system;
2. The gfortran compiler (the program was tested for version 4.8)
3. The gnuplot;
4. The R compiler;
5. The flowcore package;

2 Compiling

The first step is to download all the content in the github repository (https://github.com/CombinatorialCytometry/multiparametric-flow-cytometry). And make sure the computer has all the dependencies installed and working properly. To compile the code, open a Linux terminal in the same folder where the files were saved. Then compile the code using "make":

```
> make
```

If the compilation goes well and there is no compilation errors, you should see a file named "multiparametric-flow-cytometry". To execute the program just type into the linux terminal:

```
> ./multiparametric-flow-cytometry
```

To get the graphics with the colored dot-plots we need to provide an input configuration ("in.dat"). In the downloaded files we provide some examples that show all the data using all the fluorescent signals to calculate the color tendency.

3 The Configuration File

To build the colored dotplots it is necessary to provide additional information to the program. Here we show an example configuration file and explain how to edit it in order to perform different analysis in the data.

The input file must be named as "in.dat". One available example is show below:

```
1   Sample_39.fcs
2   10
3   6
4   1 5 6 7 9 10 11 12 14 15
5   525 625 475 675 725 425
6   10.09  17.35  10000.0  46.76  44.37  13.47
```

```
7   100.00    1.40    0.00    0.00    0.00     0.00
8    18.22  100.00   11.49    0.00    0.00     0.00
9     6.07   28.50  100.00    0.00    0.00     0.00
10    2.80   15.42   20.11  100.00    0.00     0.00
11    0.00    0.41    0.00    6.86  100.00    14.49
12    0.00    0.00    0.00    0.00    6.54   100.00
13  0 0 0 0 0 0
```

The file is composed of 13 lines. It is important to keep the correct order of the inputs. We can split the inputs in eight items, organized by it their function in the process.

```
1) name of the FCS file.........................Sample_39.fcs
2) total number of columns in the fcs file.......10
3) number of fluorescent channels................6
4) columns to copy into memory..................1 5 6 7 9 10 11 12 13 14 15
first the fsc, ssc and additional data
after the fluorescent channel
5) wavelength values for the channel colors......525 625 475 675 725 425
6) Background values............................10.09  17.35  10000.0  46.76  44.37  13.47
7) Compensation values (the full matrix).....................................
...100.00    1.40    0.00    0.00    0.00     0.00
....18.22  100.00   11.49    0.00    0.00     0.00
.....6.07   28.50  100.00    0.00    0.00     0.00
.....2.80   15.42   20.11  100.00    0.00     0.00
.....0.00    0.41    0.00    6.86  100.00    14.49
.....0.00    0.00    0.00    0.00    6.54   100.00
8) Fluorescent channels used in the color trend..0 0 0 0 0 0 0
```

The in.dat must provide these information in order, without any other text.

1) The first line is the complete name of the Flow Cytometry Standard file. The program will read this FCS file and analyze it. 2) The second line is the number of columns that the program will copy from the FCS file.* 3) Number of fluorescent signals. For example, if you want to analyze seven filter signals and five more information (time, Id, etc.) The second line must have 12 and the third line must have 7.* 4) The columns to copy, here you need to know which columns of the FCS file has the information you want to put into the color dotplot. Please put first the general data (like event ID, time, morphology, etc) and then the fluorescent columns.*

*If you do not know which columns to choose you can open the FCS file and read the first line or execute the multiparametric-flow-cytometry program. The program will show you the first line of the file with all the names of the columns for you to choose. You will need to count the position of the data. The counting star in 1 (sorry for that C programmers!!!).

5) The wavelength of each fluorescent channel in order. value must be in [400, 750]nm, the values can be related or not with the real filters values. If you choose to not use the filter half value, we suggest to use wavelength values equally spaced. Like the example.

6) The background value is the fluorescence intensity to consider the event as a positive one for the channel. This value must be in agreement with the data.

7) The traditional compensation table. The values must be in [0,100] and 100 for the channel with itself.

8) To calculate the color tendency for all the channels, let this value as zeros (one zero for each color channel) or complete with the sequence until the number of channels (e.g 1 2 3 4 5 6). This way all the channels are taken into account to the tendency color. To calculate the tendency only for events positive for just a set of channels put their numbers followed by zeros, e.g. 2 4 0 0 0 0. In this case only the double positives events for the second and fourth channels will receive a color, all the other are going to stay black. And to calculate the tendency for the negative events for a set of channels just put the number of channels in negative, e.g 2 -4 0 0 0 0. With this configuration the events are going to receive a color if and only if they are positive for the second channel and negative for the fourth channel. It is important to keep the zeros, the total number of digits must be equal to the number informed in line 2

4 Example

If you run the example described in this tutorial, you should obtain two exit files: "result_file.eps" with the colored dotplot and "result_file_deb.eps" with the trust values. The results are shown in figure 1

Please make sure you rename output files before running new tests otherwise the output files will be overwritten.

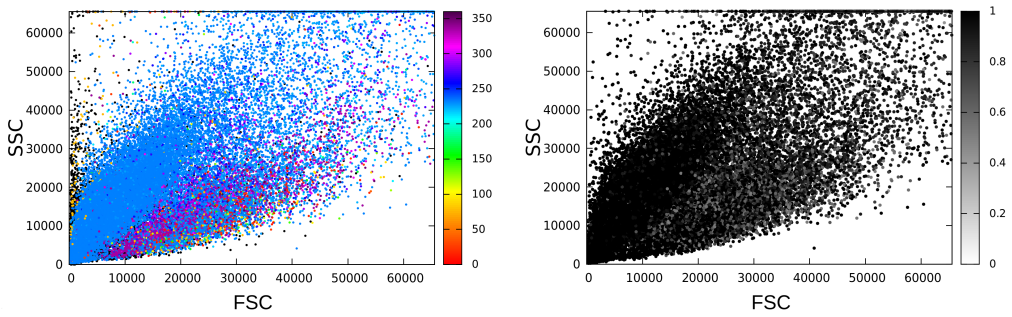

Figure 1: In the left, the colored dot-plot. In the right, the trust values (for the 6 colors).

4.1 Selection of Positive Filters

For the flow cytometry analysis it may be interesting to see not all the colors in the tendency plot at the same time. In some situations, when showing all the colors the results may have low trust or even hide populations due to the high number of other cell types in the same region. To select only events with positive signal for the channel 1 (FITC in the example case) edit the file "in.dat". Change the line number 13 from "0 0 0 0 0 0" or "1 2 3 4 5 6" to "1 0 0 0 0 0". The results for this selection are presented in figure 2.

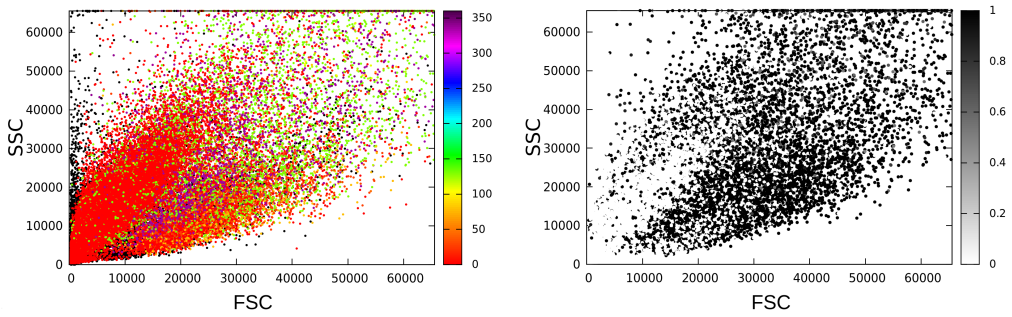

Figure 2: In the left, the colored dot-plot. In the right, the trust values. In this case we selected only the positive dots for the channel number 1 who corresponds to the FITC marker. Notice the trust plot, the white region suggests that those dots are not well defined by the color marked and they should not be taken into account.

4.2 Selection of Negative Filters

For the flow citometry analysis it may be interesting to see not all the colors in the tendency. Because with all the colors the results may have low trust or even hide populations due to the high number of other cell types in the same region. To select only events with negative signal for the channel 1 (FITC in the example case) edit the file "in.dat". Change the line number 13 from "0 0 0 0 0 0" or "1 2 3 4 5 6" to "-1 0 0 0 0 0". The results for this selection are presented in figure 3.

Comparing figures 2 and 3 we can see that the regions colored are different. In fact, all the dots colored in the first one are now black. The only dots black in both are the events negative for all the fluorescent channels.

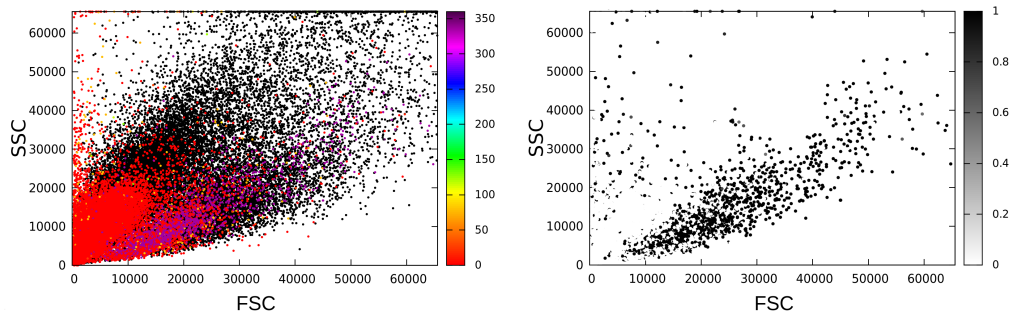

Figure 3: In the left the colored dotplot. In the right the trust values. In this case we selected only the negative dots for the channel number 1 who corresponds to the FITC marker. Notice the trust plot, the white region suggests that those dots are not well defined by the color marked and they should not be taken into account.

4.3 Selection of Filter Combinations

It is possible to select any combination of channels for positive or negative events. To select a combination of positive and negative values just edit the input file "in.dat" in line 13. For example, if you want to see the color tendency for all events with the second filter positive and the fourth filter negative, the line 13 must be like: "2 -4 0 0 0 0".

5 Gnuplot Scripts

The multiparametric-flow-cytometry creates two temporary files and the images are generated in gnuplot by the two scripts provided. The multiparametric-flow-cytometry call the gnuplot automatically. The script provided in the downloaded files will work with the configuration file. However is up to the user to edit the gnuplot scripts in any new data.

To create the colored dot-plot the script below read the temporary files and plot them into the final eps images.

```
1....set termoption enhanced
2....set title font ", 20"
3....set xlabel 'FSC'
4....set xlabel font ", 18"
5....set xtics font ", 11"
6....set ylabel font ", 18"
7....set ytics font ", 11"
8....set ylabel 'SSC' enhanced
9....set nokey
10...set cbrange [0:360]
11...set palette defined ( 0 '#ff0000',\
.......................1 '#ff7d00',\
.......................2 '#ffff00',\
.......................3 '#00ff00',\
.......................4 '#00ffff',\
.......................5 '#0000ff',\
.......................6 '#ff00ff',\
.......................7 '#440044')
12...plot [0:65533][0:65533] "nofl.dat" using 2:4:5 with points
     pt 7 ps 0.3 lc rgb "black"
13...set term postscript eps enhanced color
14...set output "result_file.eps"
15...replot "fluo.dat" using 2:4:5 with points palette pt 7 ps 0.3
```

3... Set the x axis label, the default is FSC.
4... Set the x axis label font size.

6... Set the y axis label font size.
8... Set the y axis label, the default is SSC.
11.. Define the palette, i.e. the colors of the graphic bar. The default values are an visual approximation of the wavelength values.
12...The values inside the braces are the region ploted, for example, [0:1024][0:1024] will plot only the values between 0 and 1024 in both axes.

After the word using there are tree numbers, these number are very important, they are the source of the dot-plot. The first one is the column where the FSC information is saved. The second one is where the SSC information is saved. The third one is where the color is saved.

To know where the data you want to plot are in the temporary files is a manual process.

In the configuration file, you set the number of total columns and the number of flourescent channels, the temporary files will keep the other data and put the information in the next row. So, in the example file, in line 2 we have that the total data is equals to 10 and in line 3 we have 6 fluorescent channels, so there are 4 columns of other data (such as event ID, FSC, SSC). This way the color data will be placed in the fifht column.

When you choose the colums in the configuration file, in the line number 4 you choose as well the order of the temporary files columns. If you keep the second column as the FSC and the fourth column as the SSC. There is no need to edit this script. However it will depend on how you selected the data in the configuration file.

The lines 13 and 14 set the output format and file. 15... the same number edition made in line 12 are required here.