Multiparametric Flow Cytometry First Steps

Carine P. Beatrici

January 21, 2018

1 Introduction

The multiparametric-flow-cytometry is the program 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 4.8 version)
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 install and properly working. To compile the code, go to a Linux terminal in the folder that the files were saved. Then compile the code with "make":

```
> make
```

If the compilation went well and there was no compile errors in the same folder you should see a file named "multiparametric-flow-cytometry". To execute the program just type into linux terminal:

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

To get the graphics with the colored dot-plots we need to have a well written input file ("in.dat"). In the downloaded files we provide an example that shows all the data using all the fluorescent signals to calculate the color tendency.

3 The Input File

To build the colored dotplots it is necessary to provide the informations to the program. Here we show an example 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" the example downloaded 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 by it its 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 have this informations 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 interest data, 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 filters signals and five more informations (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 this example files you should get 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 copy the results to other files before running new tests otherwise you may rewrite the file with the new test.

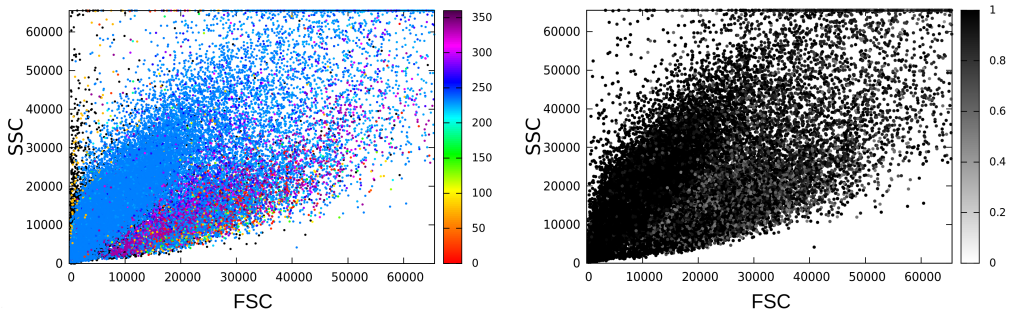

Figure 1: In the left the colored dotplot. In the right the trust values. For the 6 colors.

4.1 Selection of Positive 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 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.

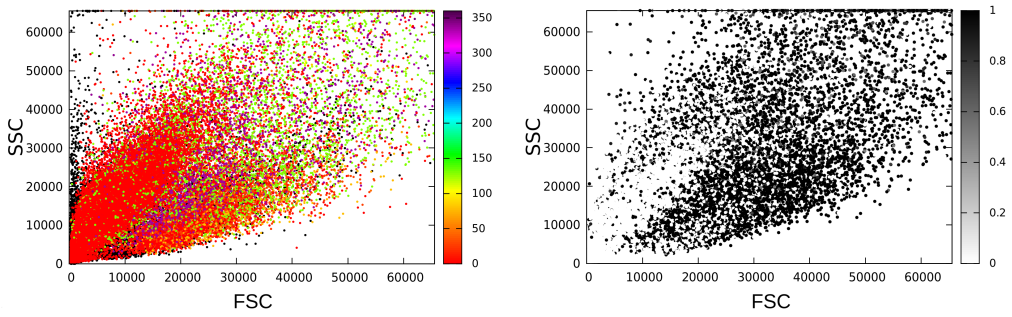

Figure 2: In the left the colored dotplot. In the right the trust values. In this case we selected only the positive dots for the channel number 1 who corresponds to the FITC marker. Notice the trust plot, the white region suggests that those dots are not well defined by the color marked and they should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

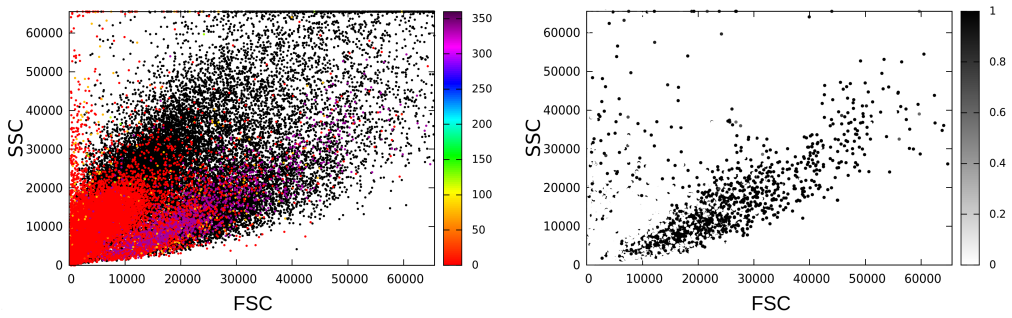

Figure 3: In the left the colored dotplot. In the right the trust values. In this case we selected only the negative dots for the channel number 1 who corresponds to the FITC marker. Notice the trust plot, the white region suggests that those dots are not well defined by the color marked and they should 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. The criteria must be biological to determine if this combination makes sense. 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".