

How to: cytometry data preprocessing and quality control in R

Sarah Bonte



R basics

script:
store code

console:
execute code

The screenshot displays the RStudio interface with the following components:

- Script Editor:** Contains R code for variable assignment and arithmetic operations.
- Environment Pane:** Shows the current environment with variables 'a' (value 5) and 'A' (value 6).
- Console:** Shows the execution output of the code, including variable values and arithmetic results.
- Environment Tab:** Displays the 'PeacoQC' package and its documentation.

```
6
7 # variables ----
8
9 # Multiple types of data in R
10 1
11 "a"
12 TRUE
13
14 # Store in variables
15 a <- 5
16 # Show what is stored in a variable (also look at environment tab)
17 a
18 # Variables are case sensitive
19 A <- 6
20 A
21 a
22 # Reuse the variable
23 a + a
24
25 # Ex: Store a number of your choice in a new variable
26 b <- 3
27 # Ex: Multiply a and b
28
```

Environment

Values	
a	5
A	6

Console

```
R 4.3.0 · D:/postdoc/workshops/MBenC course 240208/
> # Store in variables
> a <- 5
> # Show what is stored in a variable (also look at environment tab)
> a
[1] 5
> # Variables are case sensitive
> A <- 6
> A
[1] 6
> a
[1] 5
> # Reuse the variable
> a + a
[1] 10
>
```

PeacoQC (PeacoQC) R Documentation

Peak-based detection of high quality cytometry data

Description

PeacoQC will determine peaks on the channels in the flowframe. Then it will remove anomalies caused by e.g. clogs, changes in speed etc. by using an IsolationTree and/or

environment:
collection of
user-defined
objects

R basics

script:
store code

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code with comments. A white arrow points from the text "place comments after #" to the comment on line 9. Another white arrow points from the text "CTRL + ENTER to execute line" to the end of line 25.
- Environment Tab:** Shows the current environment with variables 'a' (value 5) and 'A' (value 6).
- Console:** Shows the execution output for the code entered in the source editor.
- Viewer Tab:** Displays a document titled "Peak-based detection of high quality cytometry data" with a description of the PeacoQC package.

```
6
7 # variables ----
8
9 # Multiple types of data in R
10 1
11 "a"
12 TRUE
13
14 # Store in variables
15 a <- 5
16 # Show what is stored in a variable (also look at environment tab)
17 a
18 # Variables are case sensitive
19 A <- 6
20 A
21 a
22 # Reuse the variable
23 a + a
24
25 # Ex: Store a number of your choice in a new variable
26 b <- 3
27 # Ex: Multiply a and b
28
```

Environment

Variable	Value
a	5
A	6

Console

```
R 4.3.0 · D:/postdoc/workshops/MBenC course 240208/
> # Store in variables
> a <- 5
> # Show what is stored in a variable (also look at environment tab)
> a
[1] 5
> # Variables are case sensitive
> A <- 6
> A
[1] 6
> a
[1] 5
> # Reuse the variable
> a + a
[1] 10
>
```

Viewer

Peak-based detection of high quality cytometry data

Description

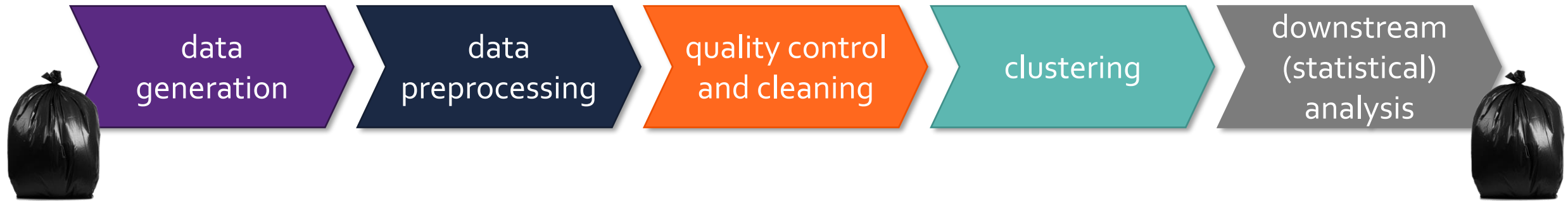
PeacoQC will determine peaks on the channels in the flowframe. Then it will remove anomalies caused by e.g. clogs, changes in speed etc. by using an IsolationTree and/or

R basics

variable, function, for loop, ...



Automated computational analysis pipeline



garbage in

=

garbage out

Automated computational analysis pipeline



FCS file



fcs file

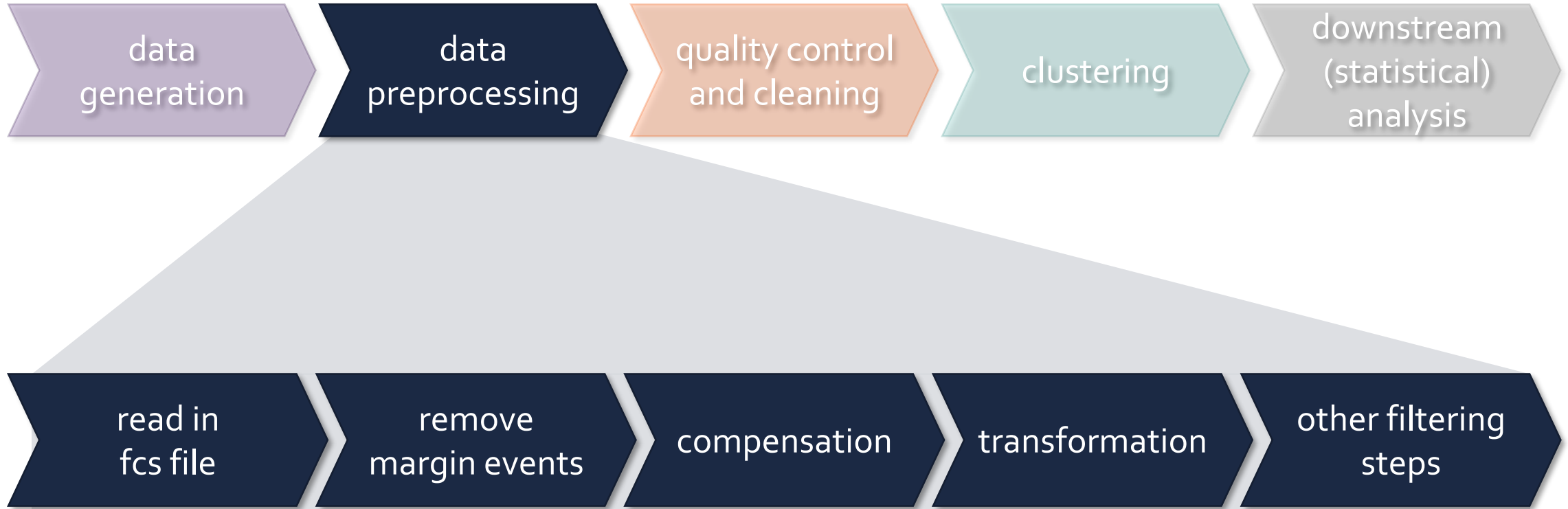
install and load libraries



set and create directories

load in fcs file

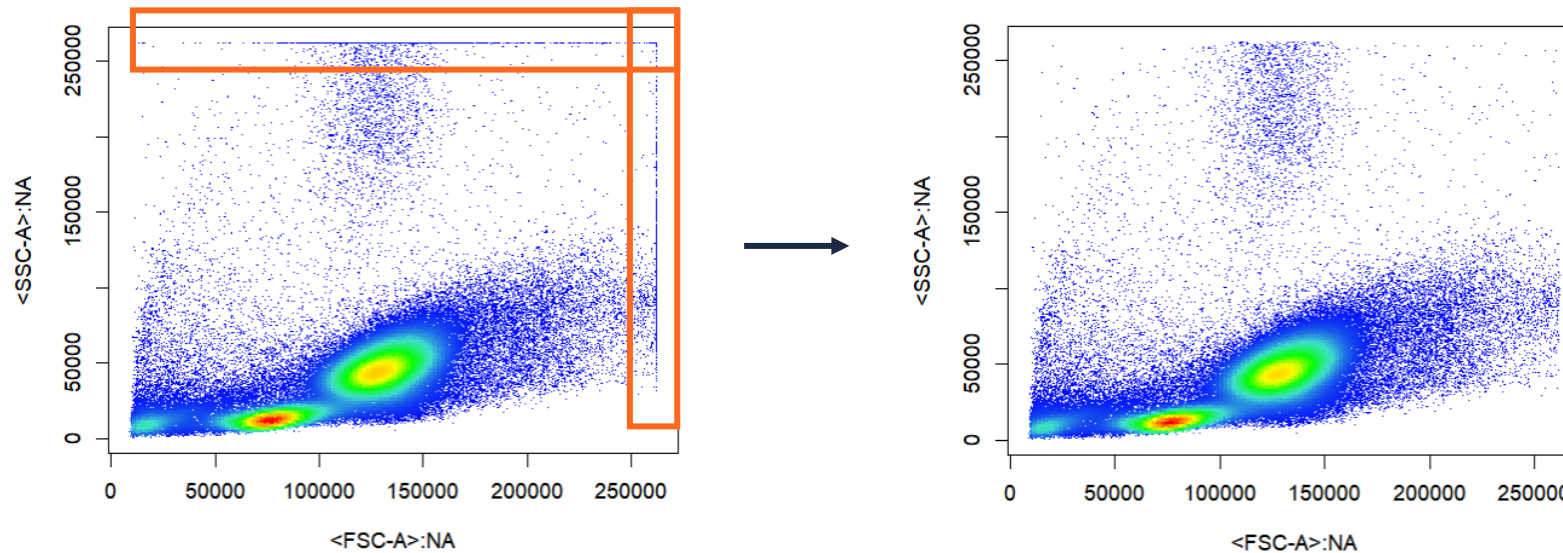
Automated computational analysis pipeline



Preprocessing overview



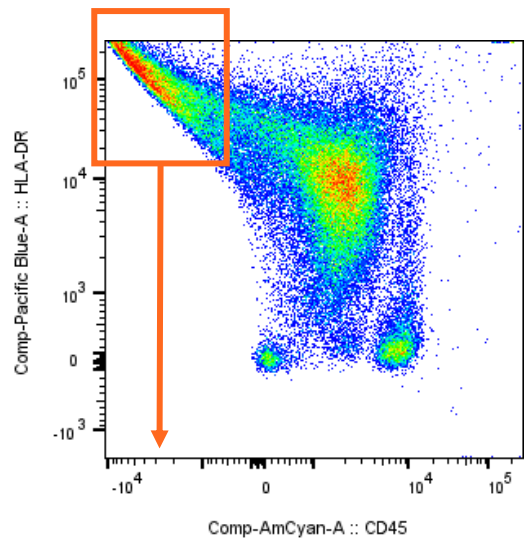
Remove events that are outside, or at the borders of, the detectable range of the cytometer



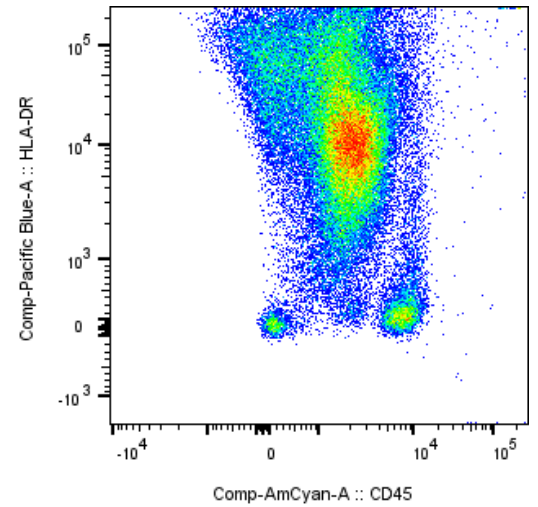
Preprocessing overview



	Pacific Blue-A	AmCyan-A
[1,]	332.28000	335.12000
[2,]	602.07996	450.13998
[3,]	859.09998	487.06000
[4,]	624.79999	320.91998
[5,]	827.85999	612.01996
[6,]	1404.38000	576.51996
	90000	- 10000

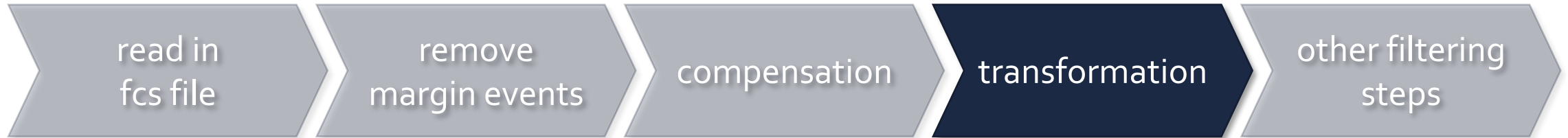


e.g. overcompensation

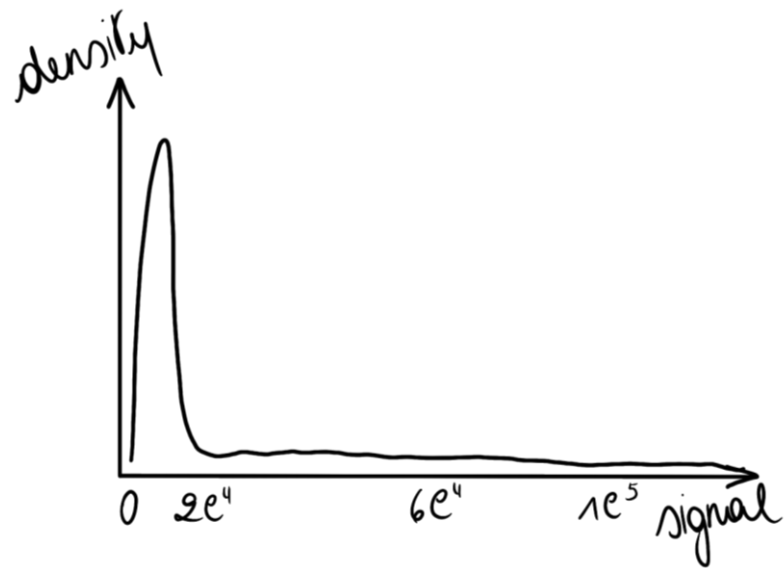


correct compensation matrix applied

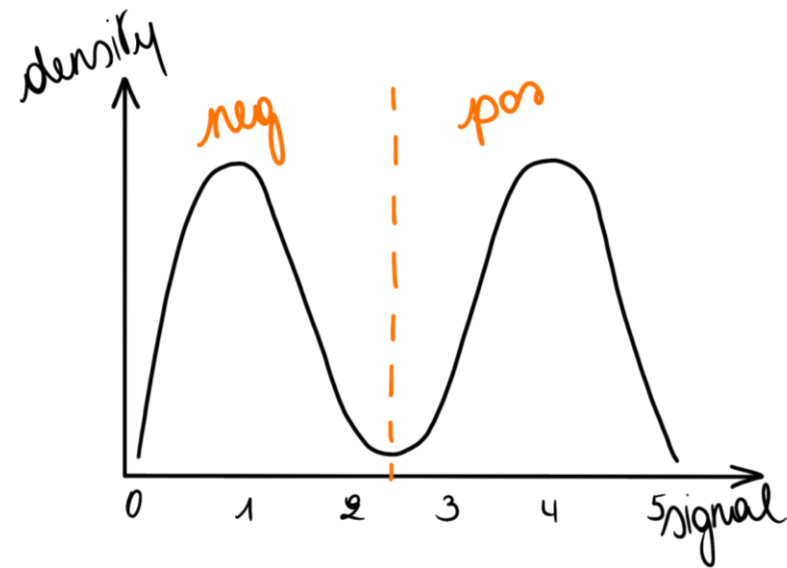
Preprocessing overview



without transformation



biexponential transformation (e.g. logicle)



Preprocessing in R

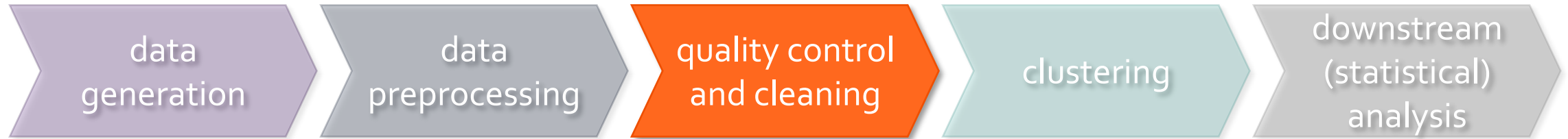
RemoveMargins



compensation

transformation

Automated computational analysis pipeline



Quality control on two levels



per file

e.g.

- clogs during acquisition
- changes in flow rate

check the signal consistency over time



between files

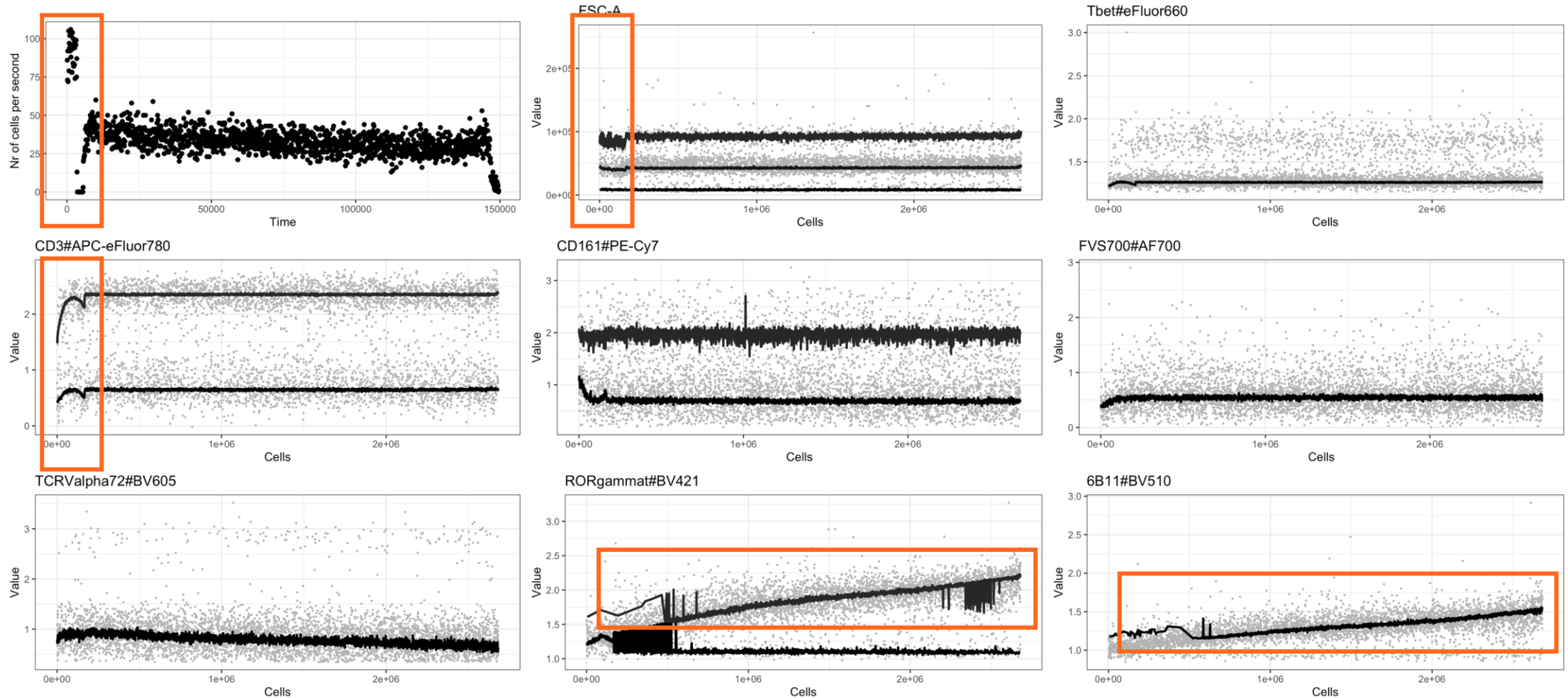
e.g.

- new antibody batch
- change in machine settings

check for batch effects between samples

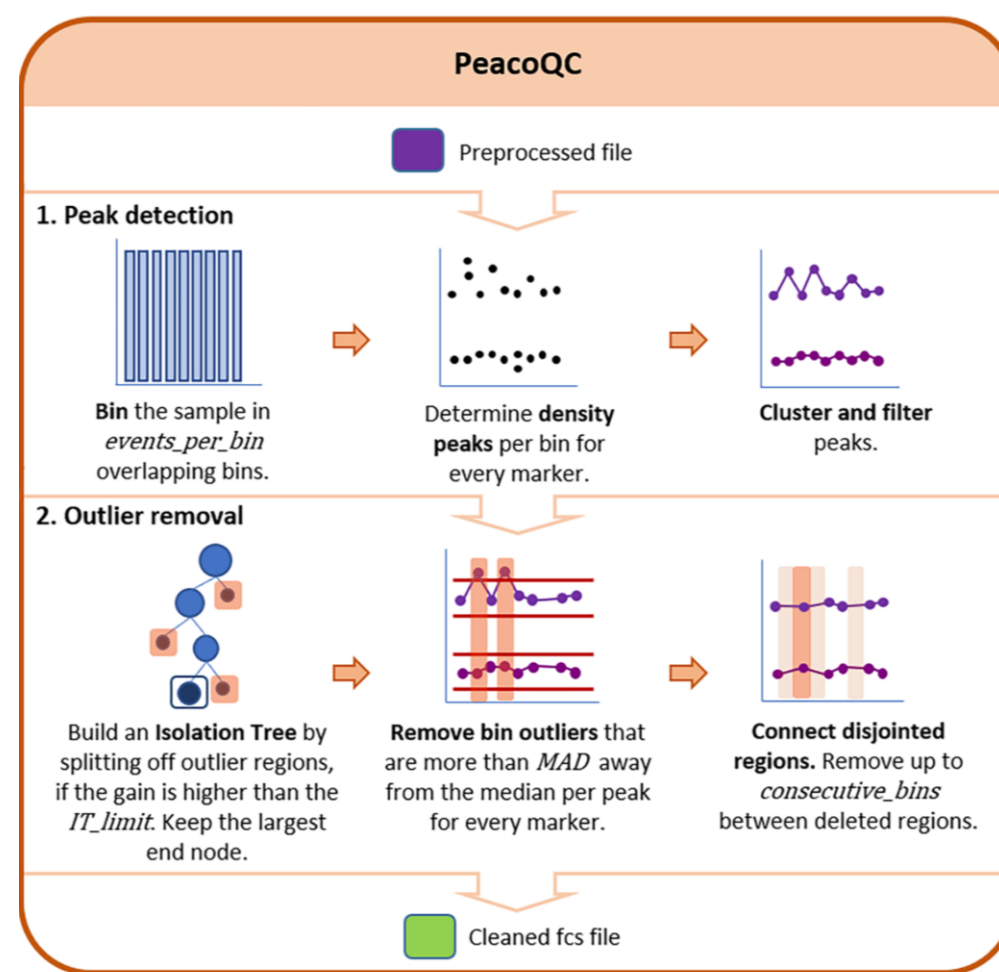
Within-file quality

changes in flow rate and signal



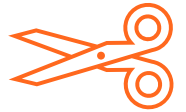
changes in signal

PeacoQC for automated quality control

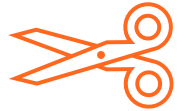


PeacoQC for automated quality control

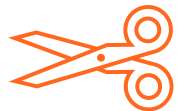
	FSC-A	FSC-H	SSC-A	B-530/30-A	B-585/42-A	B-670LP-A	B-780/60-A	R-660/20-A	R-780/60-A	V-450/50-A	V-510/50-A
[1,]	143804.703	112032	37015.1406	332.28000	335.12000	346.47998	53.96000	360.099976	-10.400000	1097.360	3913.000
[2,]	166320.891	121620	57066.9570	602.07996	450.13998	501.25998	183.17999	488.799988	197.599991	2057.980	5333.720
[3,]	193674.594	138753	72303.5547	859.09998	487.06000	536.76001	144.84000	673.399963	321.099976	1772.460	4105.640
[4,]	156393.000	122200	47006.2578	624.79999	320.91998	245.65999	51.12000	555.099976	609.699951	1177.340	4117.680



[8,]	86222.695	74144	9542.3994	153.36000	203.06000	323.75998	88.04000	5894.199707	458.899994	14871.120	8632.680
[9,]	178893.000	134069	51327.3203	746.91998	467.17999	549.53998	167.56000	527.799988	-28.599998	1755.260	3973.200
[10,]	247311.891	167984	55742.0977	880.39996	542.44000	553.79999	21.30000	821.599976	53.299999	2154.300	5132.480
[11,]	170807.391	131809	36185.8594	462.91998	399.01999	305.29999	31.24000	445.899994	208.000000	1319.240	4938.980
[12,]	182610.000	136931	65892.2578	15655.50000	2433.87988	1226.88000	181.75999	756.599976	-16.900000	1520.480	3100.300
[13,]	157763.703	109292	102329.4531	958.50000	498.41998	633.32001	79.52000	1127.099976	-52.000000	2162.040	2174.940
[14,]	157670.094	123955	54684.1992	14040.95996	1901.37988	1486.73999	299.62000	513.500000	-26.000000	1587.560	3154.480
[15,]	149195.703	116999	27406.0000	685.85999	195.95999	952.81995	143.42000	1601.599976	7.800000	1015.660	1400.080



[17,]	163785.594	117165	28897.0000	20594.25977	2597.17993	776.73999	259.85999	815.099976	-3.900000	1131.760	2724.480
[18,]	35729.098	31493	33698.0195	306.72000	122.12000	452.97998	222.93999	317.199982	209.299988	473.000	836.780
[19,]	128565.898	106740	39812.5391	17013.01953	2379.91992	1581.88000	668.82001	339.299988	1622.399902	1571.220	3267.140
[20,]	176066.094	134946	46135.7969	573.67999	400.44000	479.95999	39.76000	674.699951	422.499969	1787.080	6928.160
[21,]	161051.391	122387	45881.6172	474.28000	286.84000	292.51999	-75.25999	460.199982	401.699982	1747.520	4645.720
[22,]	164090.703	125696	29232.1191	4651.91992	843.47998	931.51996	72.42000	505.699982	39.000000	2000.360	2545.600
[23,]	154710.891	91185	26857.8789	404.69998	332.28000	592.13995	435.93997	22634.298828	19925.099609	14140.980	13489.101
[24,]	151389.891	113838	32080.6387	2335.89990	509.78000	528.23999	66.74000	293.799988	218.399994	1927.260	6317.560
[25,]	165077.094	127851	88765.6172	14587.65918	1965.27991	1500.93994	228.62000	916.499939	109.199997	1613.360	2946.360
[26,]	83719.797	63033	23221.2598	502.67999	63.90000	289.67999	-38.34000	349.699982	148.199997	1444.800	6542.880
[27,]	190896.297	142590	49665.9180	587.88000	265.53998	489.89999	113.60000	384.799988	421.199982	1605.620	4938.980



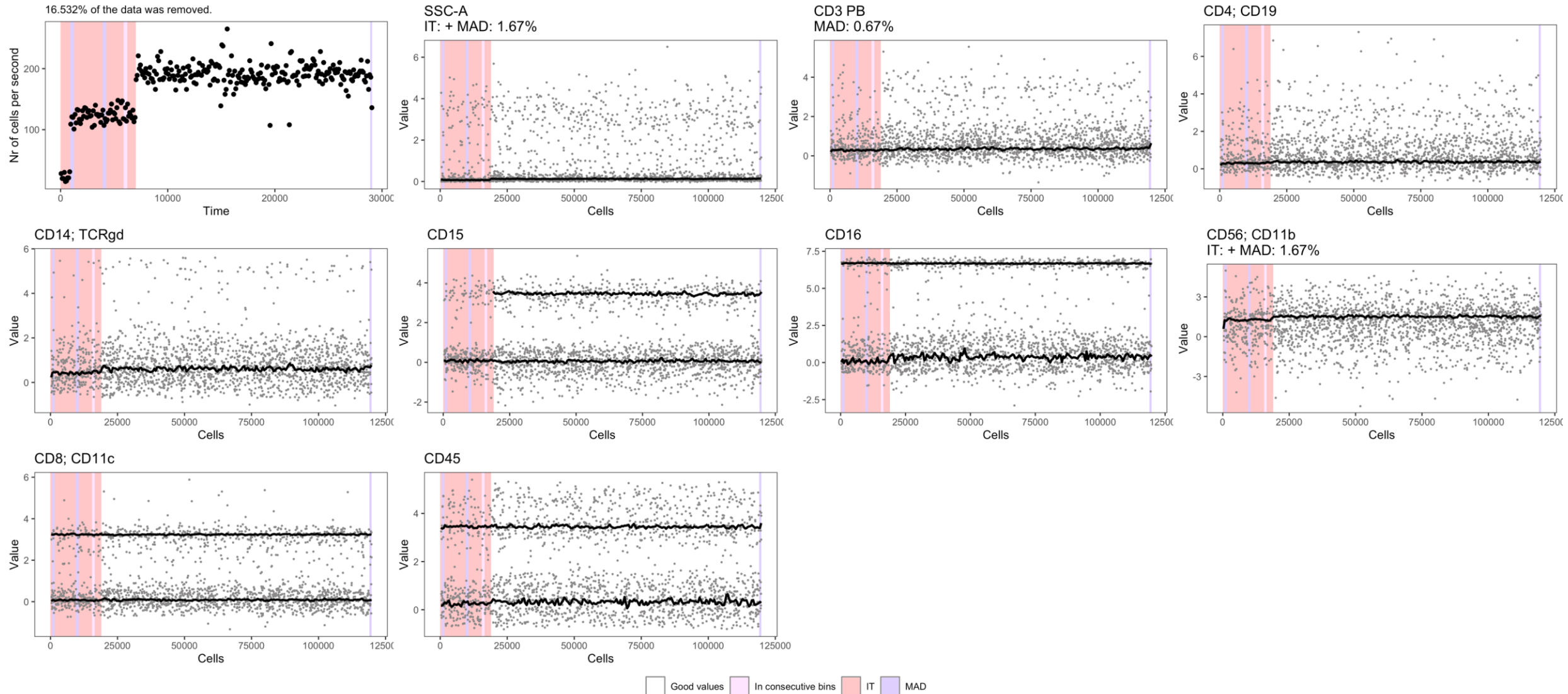
[30,]	90050.398	57656	46617.1797	850.57996	505.51999	1187.12000	619.12000	492.699982	12064.000000	1302.900	8154.520
[31,]	170381.703	128415	36090.7188	51757.57812	6458.15967	2797.39990	143.42000	436.799988	244.399994	1245.280	4130.580
[32,]	135000.891	105994	17315.4785	2093.07983	512.62000	424.57999	-9.94000	5820.099609	412.099976	963.200	1612.500
[33,]	213623.094	155964	64899.6797	626.21997	523.97998	705.73999	53.96000	530.399963	178.099991	1583.260	5214.180
[34,]	262143.000	132466	70965.9141	1811.91992	575.09998	962.75995	373.45999	1131.000000	366.599976	2670.300	8816.720
[35,]	138026.703	102190	33558.8594	816.50000	242.81999	773.89996	211.57999	568.099976	383.500000	1569.500	6247.900

PeacoQC for automated quality control



	FSC-A	FSC-H	SSC-A	B-530/30-A	B-585/42-A	B-670LP-A	B-780/60-A	R-660/20-A	R-780/60-A	V-450/50-A	V-510/50-A
[1,]	143804.703	112032	37015.1406	332.28000	335.12000	346.47998	53.96000	360.099976	-10.400000	1097.360	3913.000
[2,]	166320.891	121620	57066.9570	602.07996	450.13998	501.25998	183.17999	488.799988	197.599991	2057.980	5333.720
[3,]	193674.594	138753	72303.5547	859.09998	487.06000	536.76001	144.84000	673.399963	321.099976	1772.460	4105.640
[4,]	156393.000	122200	47006.2578	624.79999	320.91998	245.65999	51.12000	555.099976	609.699951	1177.340	4117.680
[8,]	86222.695	74144	9542.3994	153.36000	203.06000	323.75998	88.04000	5894.199707	458.899994	14871.120	8632.680
[9,]	178893.000	134069	51327.3203	746.91998	467.17999	549.53998	167.56000	527.799988	-28.599998	1755.260	3973.200
[10,]	247311.891	167984	55742.0977	880.39996	542.44000	553.79999	21.30000	821.599976	53.299999	2154.300	5132.480
[11,]	170807.391	131809	36185.8594	462.91998	399.01999	305.29999	31.24000	445.899994	208.000000	1319.240	4938.980
[12,]	182610.000	136931	65892.2578	15655.50000	2433.87988	1226.88000	181.75999	756.599976	-16.900000	1520.480	3100.300
[13,]	157763.703	109292	102329.4531	958.50000	498.41998	633.32001	79.52000	1127.099976	-52.000000	2162.040	2174.940
[14,]	157670.094	123955	54684.1992	14040.95996	1901.37988	1486.73999	299.62000	513.500000	-26.000000	1587.560	3154.480
[15,]	149195.703	116999	27406.0000	685.85999	195.95999	952.81995	143.42000	1601.599976	7.800000	1015.660	1400.080
[17,]	163785.594	117165	28897.0000	20594.25977	2597.17993	776.73999	259.85999	815.099976	-3.900000	1131.760	2724.480
[18,]	35729.098	31493	33698.0195	306.72000	122.12000	452.97998	222.93999	317.199982	209.299988	473.000	836.780
[19,]	128565.898	106740	39812.5391	17013.01953	2379.91992	1581.88000	668.82001	339.299988	1622.399902	1571.220	3267.140
[20,]	176066.094	134946	46135.7969	573.67999	400.44000	479.95999	39.76000	674.699951	422.499969	1787.080	6928.160
[21,]	161051.391	122387	45881.6172	474.28000	286.84000	292.51999	-75.25999	460.199982	401.699982	1747.520	4645.720
[22,]	164090.703	125696	29232.1191	4651.91992	843.47998	931.51996	72.42000	505.699982	39.000000	2000.360	2545.600
[23,]	154710.891	91185	26857.8789	404.69998	332.28000	592.13995	435.93997	22634.298828	19925.099609	14140.980	13489.101
[24,]	151389.891	113838	32080.6387	2335.89990	509.78000	528.23999	66.74000	293.799988	218.399994	1927.260	6317.560
[25,]	165077.094	127851	88765.6172	14587.65918	1965.27991	1500.93994	228.62000	916.499939	109.199997	1613.360	2946.360
[26,]	83719.797	63033	23221.2598	502.67999	63.90000	289.67999	-38.34000	349.699982	148.199997	1444.800	6542.880
[27,]	190896.297	142590	49665.9180	587.88000	265.53998	489.89999	113.60000	384.799988	421.199982	1605.620	4938.980
[30,]	90050.398	57656	46617.1797	850.57996	505.51999	1187.12000	619.12000	492.699982	12064.000000	1302.900	8154.520
[31,]	170381.703	128415	36090.7188	51757.57812	6458.15967	2797.39990	143.42000	436.799988	244.399994	1245.280	4130.580
[32,]	135000.891	105994	17315.4785	2093.07983	512.62000	424.57999	-9.94000	5820.099609	412.099976	963.200	1612.500
[33,]	213623.094	155964	64899.6797	626.21997	523.97998	705.73999	53.96000	530.399963	178.099991	1583.260	5214.180
[34,]	262143.000	132466	70965.9141	1811.91992	575.09998	962.75995	373.45999	1131.000000	366.599976	2670.300	8816.720
[35,]	138026.703	102190	33558.8594	816.50000	242.81999	773.89996	211.57999	568.099976	383.500000	1569.500	6247.900

PeacoQC for automated quality control

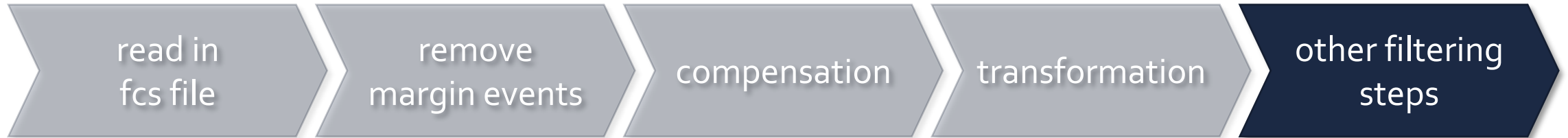


Quality control in R

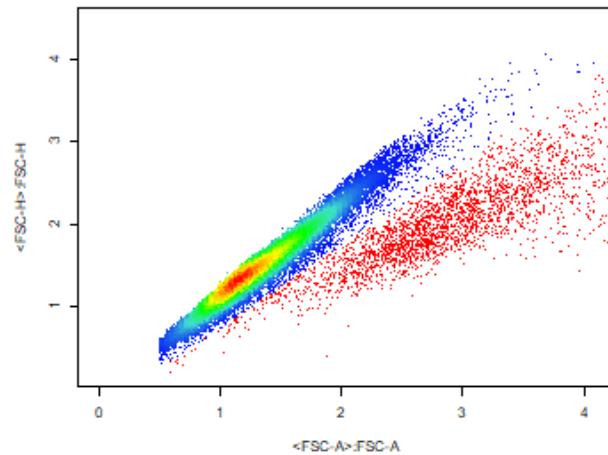
PeacoQC



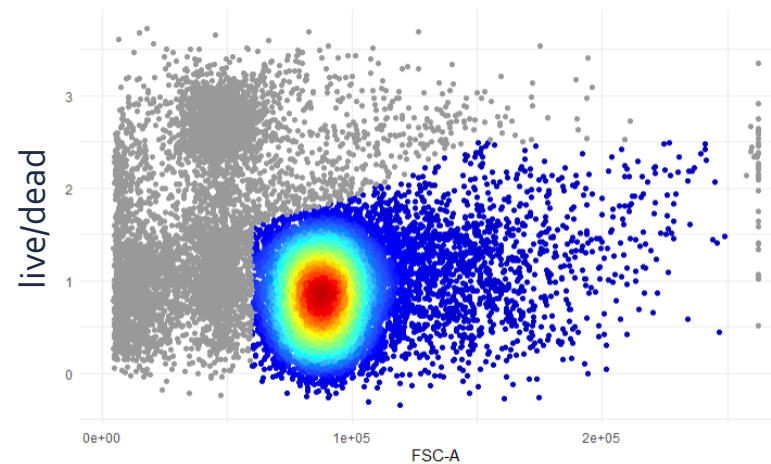
Preprocessing overview



remove doublets



remove dead cells



population of interest

Data filtering in R

RemoveDoublets



removal of dead cells (with live/dead marker)

Acknowledgements

R scripts available on
<https://github.com/saeyslab/ComputationalCytometry>

