Flow Cytometry Pipeline for Active, Dormant, and Dead Microbial Populations

M.L. Larsen

04 August, 2015

Overview

Something about microbial processes... and the type of data collected

Aims

Document workflow

Output figures, tables, and processed files

Data Collection

Data Processing

1. R Version and Package information

Source code used in this workflow were provided by X, Y University.

All analyses were completed with R version 3.2.1 (2015-06-18) and the following packages: *ADD TABLE OF PACKAGES, VERSIONS, AND DESCRIPTIONS*

1. Load data

The data for this project were collected in the IU Flow Cytometry Core Facility (C. Hassel) using eFluor fixed viability dye, Hoescht 33342, and Pyronin Y.

Due to the size, the data are batched:

1. *fs.controls* 061815_002_Hst.fcs, 061815_002_Hst_PY.fcs, 061815_002_LD.fcs, 061815_002_PY.fcs, 061815_002_UN.fcs, 061815_Ecoli_beads.fcs, 061815_Ecoli_beads_PY_Hst.fcs

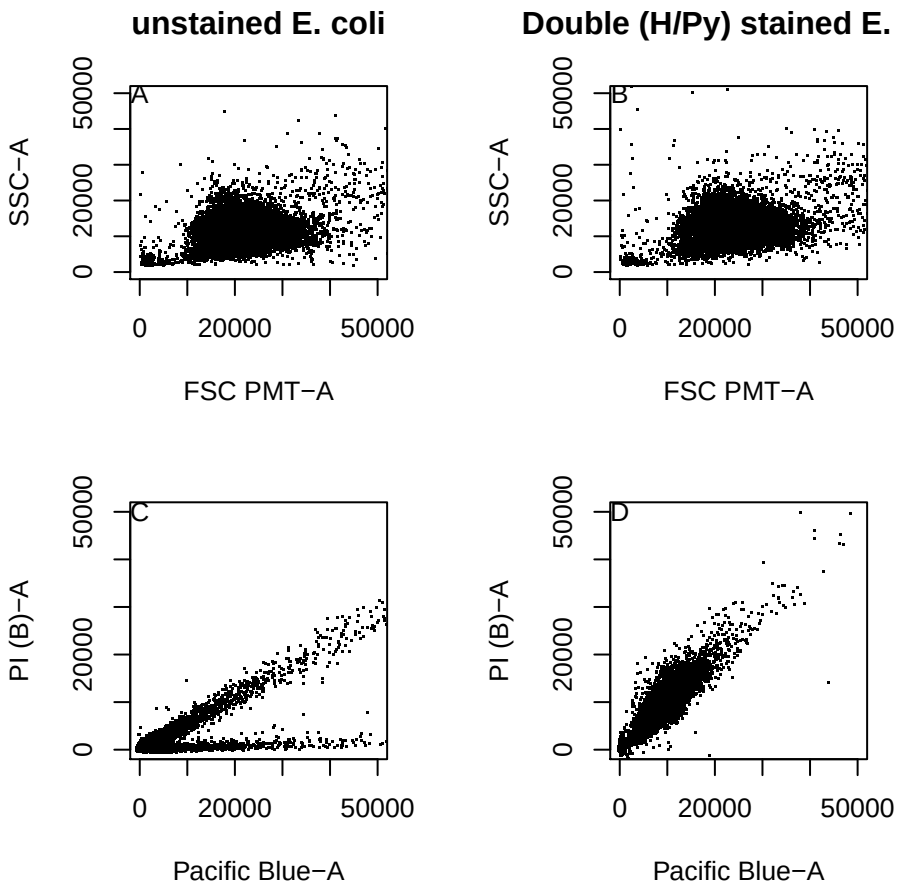
2. Diagnostics with control data

2.1 Initial Visualization

```
sampleNames(fs.controls)
length(fs.controls)

nrow(fs.controls[[1]])
fsApply(fs.controls,nrow)

fsApply(fs.controls, function(f) f@description$"TUBE NAME")

colnames(fs.controls)
```

Diagnostically evaluating the control data can help establish instrument malfunction, poor run quality, or the need for data preprocessing (**e.g.** compensations, transformation, etc.). `flowViz` provides the functionality to visually evaluate the data while `flowQ` analyzes data quality assurance.

Bivariate density plots between stained and unstained samples show shifts in the data clustering (Figure 3 B,D). The **E. coli standards** in our control batch show very similar distributions in the FSC v SSC plots, however, the distribution changes substantially with the Pacific Blue v PI plot because of the presence of the stain in the cells. We should expect a diagnol shift up in the stained group.



2.2 Preprocessing

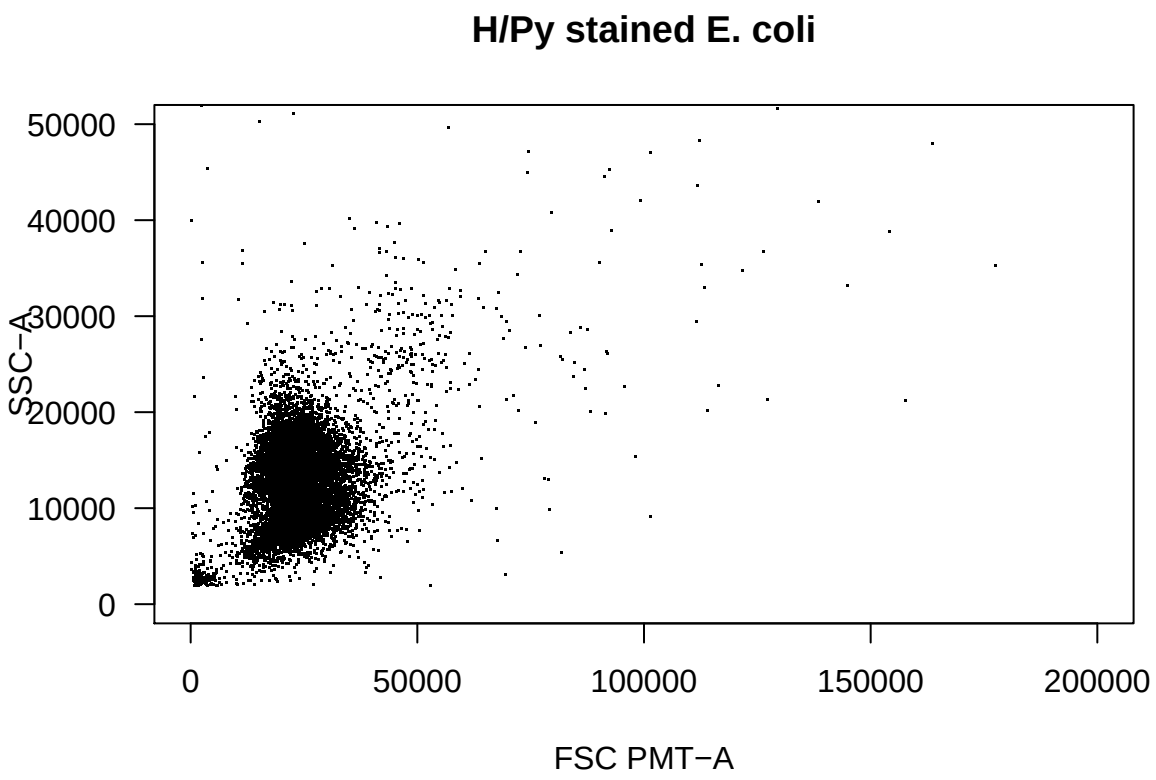
2.2.1 Compensation

Compensation is the process by which This is done because...

```
fs.controls[[1]]@description$'SPILL'

fs.controls.comp <- fsApply(fs.controls,function(frame){
  #extract compensation matrix from keywords
  comp <- keyword(frame)$`SPILL`
  new_frame <- compensate(frame,comp)
  new_frame
})

fs.controls.comp
summary(fs.controls.comp)
```

2.2.2 Visualize and Remove Margin events

Boundary, or marginal events, are common in the generation of flow cytometry data and should be considered as noise. These are events that fall out of the dynamic range that the flow cytometer can detect a signal. Before moving forward with the transformation, we must first classify and remove these boundary events with `flowQ` and **qaProcess.marginevents**, **qaProcess.cellnumber**, **qaProcess.timeline** and **qaProcess.timeflow**.



2.2.2.1 Time Anomalies **qaProcess.timeflow** & **qaProcess.timeline** can be used to detect disturbances in flow over time. Good even, flow through the machine shows that the instrument is running smoothely. These timeflow plots visualize acquisition rate over time, so the number of events that are recorded in a given time interval. The summary QA report for all the data in a **flowSet** can be found in the output.

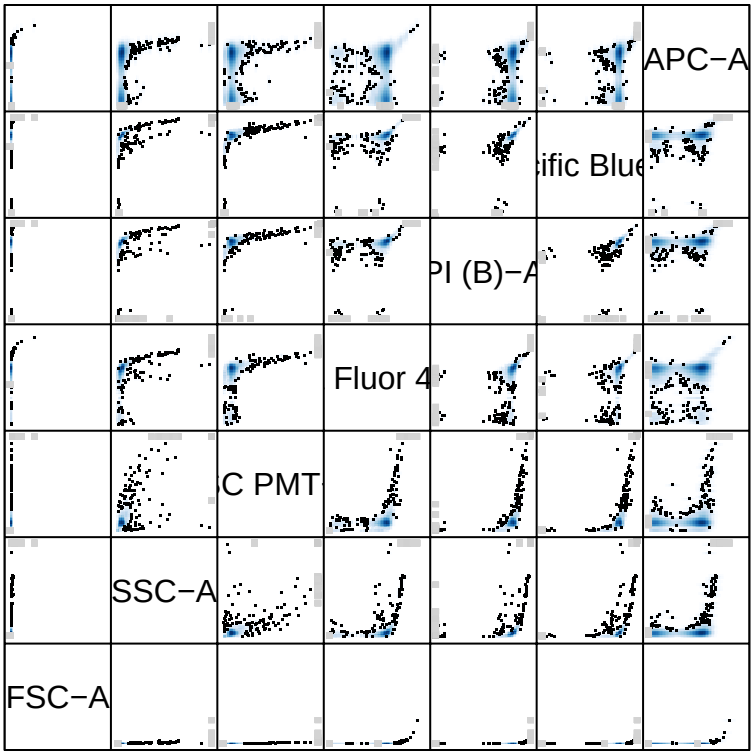
```
#qp3 <- qaProcess.timeline(fs.comtrols.comp, channels = "FSC PMT-A",
#                           outdir = dest,
#                           cutoff = 1, pdf = TRUE)
#qp4 <- qaProcess.timeflow(fs.controls.comp, channels = "FSC PMT-A",
#                           outdir = dest,
#                             cutoff = 2, pdf = TRUE)
#url <- writeQAReport(fs.controls.comp, processes = list(qp1, qp2, qp3, qp4),
#                      outdir =dest, pdf = T)
# Things look pretty good from the summary
```

Based on the QA analysis above, this flow data should be good to proceed with transformation. However, there may be some cases where the marginal events need to be removed and/or the data normalized.

2.2.3 Transformation

Data transformations can be done in-line or out-of-line techniques. Using the compenstated data, For more information about choosing the appropriate transformation, see Spidlen et al. 2006.

```
## A flowSet with 7 experiments.
##
##   column names:
##   FSC-A SSC-A FSC PMT-A Alexa Fluor 488-A PI (B)-A Pacific Blue-A APC-A
```



Scatter Plot Matrix

3. Define dynmanic gates using flowDensity

“‘{r}

Create a gate for each single, double, or triple color control as well as beads

background

beads

bead.gate <- kmeansFilter() beads.count <- summary(filter(fs[[]], bead.gate))$true

Live/Dead gating with the

livedead.gate <- kmeansFilter()

Active/Dormant gating with Hoescht 33342 and Pyronin-Y staining

“‘

4. Bring it all together now!

Active populations are cells that have RNA > DNA **Dormant** population have RNA = DNA **Dead** population is filtered out with the fixed viability dye