Automated Quadratic Characterization of Flow Cytometer Instrument Sensitivity *
(flowQB Package: Beads FCS File)

Faysal El Khettabi
Terry Fox Laboratory
British Columbia Cancer Agency
Vancouver, BC, Canada
fkhettabi@bccrc.ca
faysal.el.khettabi@gmail.com

May 26, 2014

1 Licensing

2 Introduction

We developed an automated approach to determine a cytometer's detection efficiency (Q) and background illumination (B) independent of a pre-defined bead set, based on Kmeans clustering and quadratic regression methods. We used a quadratic formulation for modelling Q and B that offers more physical insight about cytometry sensitivity than its truncated linear version [1, 3, 2], by considering the term $CV_{intrinsic}$ as part of the problem to solve. Using Kmeans in place of manual gating enabled an automated analysis to calculate Q, B and $CV_{intrinsic}$ quantities objectively and in a time-efficient manner. We validated our approach on flow cytometry sensitivity datasets through comparison to Q, B and $CV_{intrinsic}$ values obtained by manual analysis. The fully automated results ensure adequate analysis for the flow cytometry sensitivity datasets. Our approach is implemented through the R/Bioconductor package *flowQB*.

- **Methods:** We propose a collection of R generic functions to calculate Q, B and $CV_{intrinsic}$ quantities objectively and in a time-efficient manner. We

*This project was supported by the Terry Fox Foundation and the Terry Fox Research Institute and by grant 700374 from the Canadian Cancer Society, Toronto.

have implemented these functions in the Bioconductor package flowQB. We illustrate their use in this draft.

- **Results** We hope that these proposed R generic functions will become the base for the development of many tools to calculate Q, B and $CV_{intrinsic}$.
- **keywords** Flow Cytometry, High Throughput, Doublet, Instrument Sensitivity, Kmeans, Mean Fluorescence Intensity (MFI), Molecules of Equivalent Soluble Fluorochrome (MESF), linear and quadratic regressions, Q (detector efficiency) , B (background light level).

Illustration

First read the data which is in a specific folder. Our data is in flowQB-extdata folder:

RESULTS is a list

The function BEADflowQBCalculation is used for the bead FCS file to determine the singlet events. These singlet events are clustered for the channels of interest to determine the raw statitics for the regression and the generation of the regression's coefficients, Q and B values. This function generates the results as a list, the first element of the list is for Raw Statistics and the second element of the list is for the coefficients, Q and B values.

Raw Statistics uses 3 approaches, Robust Statistics, Density estimation assuming a Gaussian distribution (MASS package) and the second 'extremevalues' package used to determine the raw statistics without outliers:

- **NE:** Number of events in each peak.
- **mfiRS:** MFI associated to Robust Statistics.
- **mfiGS:** MFI associated to Gauss estimation using MASS.
- **mfino:** MFI associated to Gauss estimation using 'extremevalues'.
- **mfirSD:** Standard deviation associated to Robust Statistics.
- **mfiGS:** Standard deviation associated to Gauss estimation using MASS.
- **mfiSDno:** Standard deviation associated to Gauss estimation using 'extremevalues'.
- **Nesno.ORD:** Number of events in each peak without outliers.

For each approach (StatsProcedure) and channel (MARKER), the coefficients (c0,c1,c2) and (Q,B) are listed with their associated (Pvalue, Std-Error).

References

[1] J. Wood, *Fundamental Flow Cytometer Properties Governing Sensitivity and Resolution*, Cytometry 33, (1998), p.˜ 260 - 6.

[2] R. Hoffman and J. Wood, *Characterization of Flow Cytometer Instrument Sensitivity*, Current Protocols in Cytometry, Chapter 1: Unit 1.20 (2007).

[3] E. Chase and R. Hoffman, *Resolution of Dimly Fluorescent Particles: a Practical Measure of Fluorescence Sensitivity*, Cytometry 33 (1998), p.˜ 267-279.

[4] D. R. Parks, m. Roederer, W. A. Moore, *A new "Logicle" display method avoids deceptive effects of logarithmic scaling for low signals and compensated data.* Cytometry Part (A), (2006), 96(6), p.˜541-51.

[5] A. Ortyn, E. Hall, C. George, K. Frost, A. Basiji, J. Perry, A. Zimmerman, D. Coder, P. Morrissey, *Sensitivity measurement and compensation in spectral imaging*, Cytometry 69, (2006), p.˜ 852 - 862.

[6] *Robust Statistics in BD FACSDivaTM 6.0 Software*, BD Tech Note #23 − 9609 − 00. (2007).

[7] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, (2011),

[8] R. Ihaka and R. Gentleman R. *A Language for Data Analysis and Graphics*, Journal of Computational and Graphical Statistics, vol. 5, No. 3, (1996), p.˜299-314.

[9] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S*, Springer, Fourth Edition, New York, (2002).

[10] F. El Khettabi et al. 2014, *Automated Quadratic Characterization of Flow Cytometer Instrument Sensitivit*, to be submitted.