

FCS file conversion and transformation in ImmPort flow cytometry data analysis

2010-4-19

FCS (.fcs) files generated by instruments used in flow cytometry (FCM) experiments are in binary format and cannot be directly processed by independent analysis and visualization software. Reading and transforming a binary .fcs file requires understanding FCS file formats including FCS2.0 and FCS3.0. Unlike FCS2.0 which usually stores log-transformed data, FCS3.0 files keep the original raw outputs from the instrument in a linear mode [1]. Analysis and visualization software needs to transform the linear-mode data before delivering the data to biologists.

In our experiments, we found that most existing software that can be used for FCS file conversion, including flowCore of Bioconductor, FCSExtract, FCS2CSV, and LData, do not always provide satisfactory transformation of FCS3.0 files when compared with FlowJo MacOS version (Tree Star Inc. [2]). Figure 1 shows an example comparison between FCS2CSV and FlowJo transformations. CD3+ events are better separated from CD3- events in FlowJo-transformed data. Since FlowJo is a commercial software package, we found that it was necessary to develop an open source converter to transform binary fcs files into appropriate expression values that would be freely available to all users so that file conversion can be an integrated component for FCM analysis pipelines being developed by the bioinformatics community.

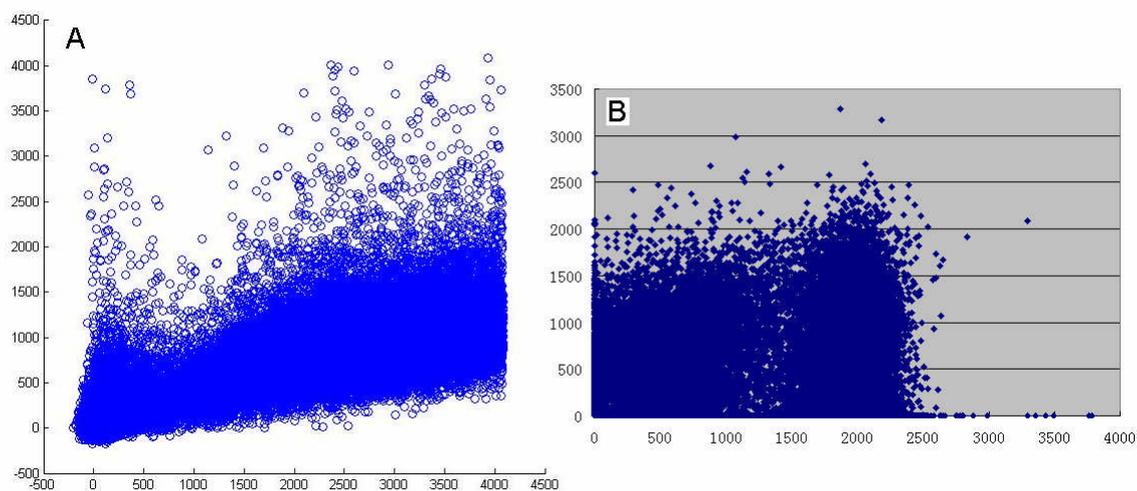


Figure 1. 2D plots of transformed values of the same .fcs file on CD3 vs. CD25 by A) FCS2CSV; B) FlowJo. CD3+ and CD3- populations are better separated in FlowJo conversion

Therefore we have developed FCSTrans, a flow cytometry data converter that generates a data matrix output from FCS2.0 and FCS3.0 binary files. The FCSTrans

data matrix output can be used by immunologists and bioinformaticians to perform lab-specific analysis and customized visualization of their FCM data. It has also been implemented in the ImmPort FCM data analysis pipeline. FCSTrans is written in R and will be made an open source Bioconductor package.

Three types of data transformations are sufficient for converting Becton, Dickinson and Company (BD) FCS2.0 and FCS3.0 files to appropriate expression values: linear, logarithmic, and logicle. While scatter data are usually transformed linearly, fluorescence channel data are traditionally transformed using logarithmic functions before being visualized or analyzed so that protein expression and other biological marker levels approximate Gaussian-like distributions, and so that different cell populations can be well distinguished. Compared with logarithmic transformation, the logicle transformation [3] usually gives rise to a better separation of populations whose expression levels are in the negative area. The logicle approach requires a careful selection of transformation parameters [4]. As the example conversion in Figure 2 shows, FCSTrans selects its transformation parameters in a way that it can generate highly consistent outputs with FlowJo.

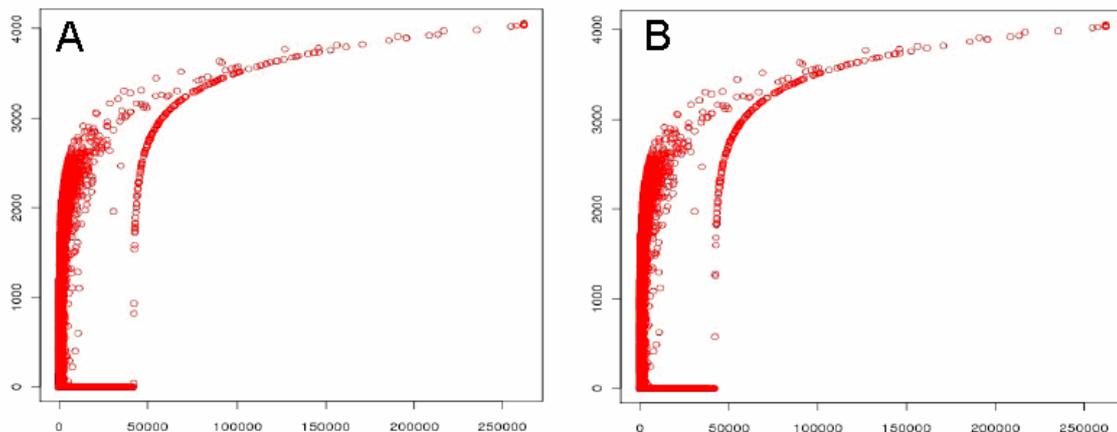


Figure 2. Plot of logicle transformed results in A) FCSTrans and B) FlowJo. Data has been compensated before the transformation. We can see the outputs of FCSTrans and FlowJo are highly consistent.

FlowTran has been tested with different types of Becton, Dickinson (BD) FCS files. The transformed results are highly consistent with those from FlowJo, which guarantees ImmPort users will see consistent dot plots or other visualizations across different software platforms. FCSTrans-converted data matrices can be downloaded from the ImmPort FCM data analysis pipeline. The ongoing conversion tests include converting files from other vendors and manufacturers and files that have been modified (gated) and re-saved after acquisition.



Acknowledgments

We sincerely thank Josef Spidlen and Ryan Brinkman for providing FCS2CSV, and the useful discussion with Florian Hahne on flowCore. This work is supported by NIH N01AI40076 (BISC).

References:

[1] BD FACS Diva Software 6.0 Reference Manual, Becton, Dickinson.

[2] <http://www.FlowJo.com>

[3] James W. Tung, Kartoosh Heydari, Rabin Tirouvanziam, Bitu Sahaf, David R. Parks, Leonard, A. Herzenberg, and Leonore A. Herzenberg. Modern Flow Cytometry: A Practical Approach, *Clin Lab Med.* 2007 September; 27(3): 453–v

[4] David R. Parks, Mario Roederer, and Wayne A. Moore. A New Logicle Display Method Avoids Deceptive Effects of Logarithmic Scaling for Low Signals and Compensated Data, *Cytometry Part A* 2006; 69A:541-551.