

BIOINFORMATICS DATA STANDARDS FOR FLOW CYTOMETRY

Ryan R. Brinkman¹, Josef Spidlen², Adam S. Treister³, Clayton Smith⁴, Perry Haaland⁵, Michael B. Ochs⁶, Robert Gentleman⁷

¹British Columbia Cancer Research Centre, Vancouver, British Columbia, Canada;

²British Columbia Cancer Research Centre, Terry Fox Laboratory, Vancouver, British Columbia, Canada;

³Tree Star, Inc., Ashland, Oregon;

⁴Terry Fox Laboratory, BC Cancer Research Centre, Vancouver, British Columbia, Canada;

⁵BD ViaSante, Carrboro, North Carolina;

⁶University of Colorado Health Sciences Center, Anesthesiology, Philadelphia, Colorado; ⁷Seattle, Washington.

Abstract. We are developing free, open source bioinformatics standards of software tools for flow cytometry (FCM) to permit a standardized approach to capturing, analyzing, and disseminating FCM data. The scope of this project includes development of an ontology, a data base schema, statistical and visual analysis tools, and a gating standard developed in the Extensible Markup Language (XML). The gating standard is presently our most well-developed component. Gating in flow cytometry is a process for defining a boundary that delineates the characteristics of particles for further analysis or sorting. During gating users define regions around a group of events of interest and then choose to include or exclude these events using some form of Boolean logic. Although flow cytometry has a successful data format standard, there is no shared representation of gates. This prevents a variety of collaborative opportunities to recreate experimental methods and results. We have developed a proposal on how to form gate definitions that can facilitate the interchange and validation of data between different software packages, and provide the means to better integrate methods with results in reporting. Specifically, we have developed a detailed description of the gating specification, a W3C Schema usable to validate XML documents, user documentation, and a set of examples of gating XML files. We have also developed an example of a software tool that implements the specification. The software tool can read an FCS data file, an accompanying XML file which describes gates according to the specification we have developed and process this information to provide descriptive statistics. Robustness problems are especially serious in geometric computation, since the numerical errors can propagate into the combinatorial computations and result in complete failure of the algorithm. We have used algorithms which are based on a user-definable precision model and contain code for robust geometric computation. As a result we feel that the software tool can be used to validate compliance of alternative software tools that implement the gating standard. The gating specification currently supports rectangular gates in n dimensions (e.g., range gates, rectangle gates, 3D box regions, and n -dimensional hyper-rectangular regions), polygon gates in n dimensions, ellipsoid gates, decision tree structures and Boolean collections of the any of the types of gates. Platform independent validation is currently implemented for 2D gates based on the JTS Topology Suite, and is available for higher dimensions in platform specific implementations based on Qhull.