Cambridge, August 11th, 2016

Dear Nature Methods editors,

We wish to submit our manuscript, “Testing for differential abundance in mass cytometry data”, for consideration as a research paper in Nature Methods.

Mass cytometry is a recently developed technique that allows researchers to simultaneously quantify the expression of many protein markers in each of millions of cells. By being able to assay more markers than conventional flow cytometry, mass cytometry provides greater resolution to distinguish between cell types and subpopulations. It is becoming increasingly used in fields such as immunology, haematopoietic development and cancer to profile complex samples. The dimensionality and complexity of mass cytometry data requires customized computational methods for analysis, and existing methods such as SPADE and X-shift have focused on identifying known and novel subpopulations by clustering cells based on their marker intensities. However, they are less suited to comparative studies where the aim is to identify differences between samples in different groups.

To this end, an alternative analytical approach is to identify “differentially abundant” (DA) subpopulations, i.e., those that change in abundance between groups. These subpopulations are interesting as they can provide some insights into the biological differences between groups. In this article, we present a novel computational method for detecting DA subpopulations in mass cytometry data. Briefly, cells from multiple samples are assigned to hyperspheres in the high-dimensional marker space, and counts for each hypersphere are tested for significant differences between groups. To correct for multiple testing, we present a method to control the spatial false discovery rate. Significant hyperspheres are then visualized in lower dimensions. This allows detection of DA subpopulations while avoiding the uncertainties and errors associated with clustering. We demonstrate the use of our method on a public data set where we are able to recover both known and previously uncharacterised subpopulations.

In summary, the method presented in this manuscript will facilitate comparative analyses of mass cytometry data in a statistically rigorous manner. We believe that the results of the DA analyses will provide greater biological insights into the differences between complex samples. We anticipate that this will be of great importance for future experiments based on mass cytometry.

We hope that our paper will be of interest to you and to Nature Methods. For reviewers, we suggest that Raphael Gottardo (Fred Hutchinson Cancer Research Center, [rgottard@fhcrc.org](mailto:rgottard@fhcrc.org)), Dana Pe'er (Columbia University, [dpeer@biology.columbia.edu](mailto:dpeer@biology.columbia.edu)), Arup Chakraborty (MIT, [arupc@mit.edu](mailto:arupc@mit.edu)) and Karen Sachs (Stanford University, [karensachs@stanford.edu](mailto:karensachs@stanford.edu)) be considered.

Yours sincerely,



John Marioni