Stony Brook University

November 9, 2018

Dr. Loyal Goff
Assistant Professor of Neuroscience
McKusick-Nathans Institute of Genetic Medicine
Johns Hopkins University School of Medicine

Dear Loyal,

I am writing to express my intent and commitment to collaborate with you and the other co-PIs on our proposal “Practical search and analysis with low-dimensional representations of the HCA“, as is being submitted in response to the *Chan Zuckerberg Institute Seed Networks for the Human Cell Atlas* request for applications. As single-cell RNA-seq biotechnology continues to progress, and more data is gathered for the HCA initiative, the development of new computational methodology, and efficient tools implementing these methods, will be critical in enabling biological discovery, and drawing understanding and insight from experimental output. I am enthusiastic about our proposed project, which will develop low-dimensional representations from single-cell RNA-seq data that can then be used to enable expression and cell profile search over the large database of HCA data, and which will allow comparison of new individual samples in the context of the HCA’s growing expression atlas.

I am committed to helping develop the base enabling technology of new accurate and computationally-efficient methods for gene quantification from single-cell RNA-seq data. Unlike existing approaches, these new methods will account for the reads (and the associated UMIs) that map between multiple genes, and will assess and propagate quantification ambiguity to our downstream tools. We will also collaborate in developing the algorithms and data structures that will enable efficient and scalable search of new experiments and user-provided expression profiles over the existing processed database. We propose to index the UMI-resolution graphs that arise from initial processing of droplet based scRNA-seq protocols, and that are the reduced data representation upon which gene-level quantification estimates are eventually computed (Aim 1). In addition to helping to lead the development of these components, I am also enthusiastic about collaborating with the other members of our network, and to provide support wherever my expertise is applicable.

I believe that the work proposed in our project will enable critical capabilities that will make the HCA data considerably more valuable, and that will help to enable and accelerate new biological discoveries. I am also very enthusiastic about our network, which includes talented and pioneering young scientists with well-aligned interests and largely-complementary skill sets. I believe this project will be a tremendous success.

Sincerely,

Rob Patro, Ph.D.
Assistant Professor of Computer Science
Stony Brook University