



SCHOOL OF PUBLIC AND ENVIRONMENTAL AFFAIRS

INDIANA UNIVERSITY

November 9, 2018

Thomas H. Hampton, Ph.D.
Senior Bioinformatics Analyst
Geisel School of Medicine at Dartmouth
Department of Microbiology and Immunology
Hanover, NH 03755

Dear Tom,

I am excited to learn of your proposal "Practical search and analysis with low-dimensional representations of the Human Cell Atlas (HCA)" that you are submitting to the Chan Zuckerberg initiative to seed networks for the HCA. My enthusiasm stems not only because of the potential significance of the HCA to humanity, but also because of my interests in training data scientists for biomedical research.

As you know from our ten years of collaborative work on gene expression analyses, I am a toxicologist at Indiana University (IU) working to characterize the molecular machinery underlying responses to chemical exposure. As I started to work with these increasingly complex data sets, I realized that most graduate and undergraduate curricula were not providing appropriate training in data science, and started to incorporate bioinformatics skills into my courses at IU and, with you and others, helped developed the Summer Workshop in Environmental Genomics that is now taught annually at the Mount Desert Island Biological lab, Maine, and the University of Birmingham, UK. Collectively the workshop has trained over 500 participants that now form a connected network of data scientists.

The HCA holds tremendous potential to uncover the mysteries of how cells work and how they are derailed by disease. However, the HCA faces the same problems encountered in the early days of genome science in that there will be a natural lag between its development and having a community of scientists with appropriate data analysis skills to mine it. Initiatives such as the one you propose, which will offer bioinformatics training in the single cell RNA-seq methods needed to contribute to and benefit from the HCA, offer hope that this lag will be minimal. The course materials you will develop will be invaluable to many and will help accelerate the pace of training. I already plan to use them in my two graduate courses in toxicology at IU and will incorporate them into the Environmental Genomics workshop.

I wish you the best in your efforts to move this initiative forward. Please let me know there is anything I can do to help.

Sincerely,

Joseph R. Shaw, Ph.D.

Associate Professor
The School of Public and Environmental Affairs
The School of Public Health
The Center for Genomics and Bioinformatics
Indiana University