Vision for postdoctoral training:
An opportunity to gain exposure to computational tools for data at scale

Word count: 791

Entering UC San Francisco six years ago was a major change from my undergraduate path. But the opportunity to blend my favorite aspects of electrical engineering with the less familiar field of biology is exactly what excited me. Starting at a biomedical research institute as an electrical engineer was a challenge. But when I learned what type of questions caught my attention and realized how to mix my interest in signal processing with the expertises of surrounding University of California at San Francisco (UCSF) environment, my research began to take off. Coming into UCSF with an unconventional background has become a strength and has led me to interacting with a diverse set of people in the community.

My desire to move from Electrical Engineering to biology wasn't random. As an undergraduate I was exposed to different forms of how engineering and biology could be mixed. It started with iGEM, a synthetic biology competition and research in a computational biology lab. While I was unable to identify what aspects of biology I liked at the time, I could sense a tremendous amount of energy and excitement at the interface of the two disciplines.

Today, I feel that there is similar opportunity in a specific combination of engineering and biology. Sequencing costs decrease faster than Moore's Law and scRNA-seq costs drop 10x per year and biologists are quickly adopting it. Two years after I entered graduate school, one of the first single cell RNA-seq datasets with just 18 cells was published, now due to automation and microfluidics, it is up 1.3 million cells. This massive growth in data will yield new insights, once we know how to best utilize it.

In order to extract biological understanding from this new scale of data, we face new challenges. We need new analyses to make sense of these complex data types. Because of their complexity, these analyses will need to be designed by people who span the computational sciences and understand the details of the measurement and biological application. For my postdoc experience I would like to immerse myself in a computer science context in order to gain exposure to new computational approaches.

I view my postdoc as an opportunity to learn a new computational techniques and tools that I can apply to solving the biological problems that I find most interesting. I will look for a research group that develops new techniques in machine learning, signal processing, and optimization designed for distributed computing platforms because these general areas have a lot of applicability to biological problems. These skills will enable me to scale scRNA-seq analyses to study the immune system and fully capitalize on the powerful scRNA-seq measurement. By working directly with the researchers developing these new techniques I can become an early adopter of their latest work and develop variations that will apply to the biological field.

While I plan to immerse myself in a research group that works on general computational problems rather than applied to a specific field, I will also evaluate the biological research groups in the wider community. Centers like the RISELab at UC Berkeley and Data Science

Institute at Columbia University have both strong computational research groups and research programs in systems biology, genomics, and immunology. I will look to integrate into these communities during my postdoc to make sure that my own work stays at the cutting edge, learn about the biological questions that they are working on, and identify the newest problems that may be able to address.

The James S. McDonnell Foundation Fellowship will give me the ability to find a research group that normally wouldn't focus on biological topics or hire postdocs without formal computer science training. The financial independence will allow to me stay focused on immunology, an area that holds great promise for improving the lives of people, while also learning computational techniques alongside the computer scientists who develop them. Additionally, any fellowship funds left over after paying for my stipend and benefits, could be applied to in collaborative data collection efforts that my new advisor may not be able to fund with his/her existing grants.

Ultimately, my long term vision is to become an academic principal investigator in a bioengineering, or computational sciences department. I will bring a computational expertise and apply it to basic biomedical research. My undergraduate degree in electrical engineering, PhD in bioinformatics at a life sciences institution, and postdoc in machine learning and techniques for big data analysis will position me to be at the cutting edge of using biological measurements as they grow to scale. I will use them to answer basic science questions about how tissues' organisation orchestrates distinct cellular populations to perform complex functionalities and how they adapt or are evaded in disease.