

In the midst of the COVID-19 pandemic, data scientists needed to effectively communicate the severity of the situation to the world. They distilled vast quantities of information into simple yet powerful visualizations, which informed pandemic response for everyone from health professionals to people without a biology background. As an aspiring professor in a quantitative discipline, one of my main goals is to make data accessible in impactful ways. I am especially curious about the intersection of biology and data science, specifically as it relates to genetics. As the volume of genetic data expands, bridging the gap between experimentalists and computationalists becomes increasingly important to advance biomedical research. Earning my PhD through the Bioinformatics program at Boston University would allow me to take the next step toward my goal of becoming a professor.

I learned to apply computational approaches to biological data as an undergraduate at Brown University, under the guidance of Dr. Neil Sarkar. In my first project, I organized data from the FDA Adverse Event Reporting System (FAERS) to pair drugs with their recorded adverse effects in an accessible reference for clinical researchers. I overcame the challenge of extracting relevant details from large data sets using MySQL and Julia. To explore my interest in molecular biology, I conducted an independent project to analyze phylogenetic relationships based on non-coding regions to understand their potential functional significance. I collated and parsed data from the NCBI Genomes database and presented these methods at a university-wide summer research symposium. These experiences provided a sound foundation in both computation and scientific communication and I became fascinated by how the analysis of biological data can provide insight into issues affecting human health.

After graduating, I joined the Genetic Perturbation Platform (GPP) at the Broad Institute of MIT and Harvard as a computational associate. I was drawn to the far-reaching impact of the group in tackling a wide range of research questions, often in collaboration with labs around the world. The GPP works in functional genomics, meaning we aim to understand the role of various parts of mammalian genomes. In the research and development group led by John

Doench, we use and optimize various technologies, such as CRISPR and its different modalities, to edit a region of the genome and measure the effect of the alteration on cells.

In one of my first projects in GPP, I identified host factors involved in SARS-CoV-2 infection by performing a meta-analysis of twelve genome-wide CRISPR screens, in collaboration with virologist Caroline Goujon and colleagues. Since each of these screens used different metrics, I could not directly compare their results in their published form. After collating raw data from each lab, I developed a standardized pipeline which I used to analyze and interpret both existing and new data sets. Compiling such a large amount of data and preparing them for integrative analysis was a significant challenge, but ultimately a rewarding learning experience. Since all of my interactions with my co-workers and collaborators were completely virtual, communicating each step of the analysis and visualizing results clearly was even more crucial. This collaboration culminated in a publication in *Nature Genetics* (PMID: 35879413), and more importantly, taught me the importance of perseverance.

In order to expand the accessibility of CRISPR screening, I worked with other members of GPP to build Apron, a web tool that allows biologists to analyze their own screening results. Our objective was to make a tool that is easy to use, can be adapted to different types of data and pipelines, and gives users agency over decisions at each step of the process. Throughout the development of the tool, I grew to understand the priorities and motivations of everyone involved in Apron's long-term success. I incorporated feedback from software engineers, platform leadership, and biologists, and formally presented the tool to the Broad community and collaborators after its beta release. This pivotal growth opportunity shaped the way I think about software development and how I value scientific communication. We continue to receive positive responses from users and believe that this tool will remain valuable to the screening community.

My interest in effective data visualization and development of accessible tools stems from my passion for teaching. Beyond my quantitative analysis experience, I have initiated computational workshops and created modules for my colleagues to discuss resources for

analyzing data. As a Broad Campus Ambassador, I speak with undergraduate students at career fairs at colleges around the country to share my enthusiasm for science and to discuss opportunities at the Broad.

I have been teaching in some capacity since I started tutoring when I was 13. Most recently, I volunteered to tutor high school students in Cambridge in a range of subjects. Recently, I joined the MITx Biology team developing online coursework for the edX platform, which offers courses to people around the world. This experience introduced me to the process of developing course materials in an inclusive manner, and I look forward to practicing and honing these skills as a teaching associate.

The Bioinformatics PhD program at Boston University will propel my journey to both make data accessible to biologists and eventually mentor the next generation of scientists. I am particularly interested in the work of Ana Fiszbein, Ruben Dries, and Joshua Campbell. While their biological areas of investigation are distinct, spanning gene regulation and cancer biology, they all share my passion for expanding the reach and impact of their research through the development of tools and software for the scientific community. Beyond my research aspirations, the Teaching Fellow requirement provides a key opportunity to develop as an educator. As an early career researcher, I have devoted my time to improving my computational knowledge, collaborating across disciplines, and learning to communicate with diverse audiences. I hope to further develop these skills as a graduate student at BU.