

The colonies were so clear that I didn't need a microscope: clusters of human melanoma cells, growing robustly despite the small molecule that should have quickly killed them. Each cell contained a guide RNA which, when paired with a CRISPR base editor, had introduced specific mutations in the drug target gene, some of which had clearly conferred drug resistance. To identify the mutations, I had to expand the colonies, extract genomic DNA, and PCR amplify and sequence the enriched guides; this process would take several weeks. Yet, looking at the flask, even my internal skeptic couldn't deny that something was happening.

This experiment was part of a project that has been my main focus as a research associate in Dr. John Doench's group in the Genetic Perturbation Platform (GPP) at the Broad Institute. I came to the GPP in 2017, after graduating with a double major in Molecular, Cellular, and Developmental Biology and Women's, Gender, and Sexuality Studies from Yale. My undergraduate research experiences -- a summer in Dr. George Church's lab at Harvard Medical School and several semesters of part-time research in Dr. Elena Ratner's lab at Yale Medical School -- had convinced me that I enjoyed being in the lab. However, both experiences were relatively supervised, and I entered my senior year unsure if a career in science was right for me. I decided that I should conduct more extensive, self-driven research before embarking on graduate study.

I was fortunate to land at the GPP, a technology platform that develops and shares tools for genetic screening in order to enable researchers across disciplines and institutions to conduct biologically meaningful and technically excellent screens. The tools we develop are widely used: our genome-wide libraries for CRISPR knockout, activation, and interference have been requested on Addgene over 2,000 times. Upon arrival, I was immediately impressed by the power of pooled screening: by pooling perturbations (in this case, guide RNAs) together in a library, we could perturb all human genes in a single screen, without the need for hundreds of 384-well plates. Still, genome-wide pooled screens aren't trivial endeavors, and I had to adapt to doing experiments on a very large scale. Whereas in my undergraduate labs I had never needed to grow more than a few flasks of cells at once, I learned to juggle several genome-wide screens, each requiring tens of millions of cells for each drug condition and replicate.

Once I adjusted to this scale of tissue culture, I realized that I would need computational skills to match. Therefore, I enrolled in an introductory computer science course in the evening at Harvard Extension School, which equipped me to analyze screening data in Python. Not only did I greatly reduce the time spent loading massive Excel files, but also I was able to more fully analyze the data that I generated. By completing the circle of an experiment -- design the library, execute the screen, analyze the data, then design the next one -- my work turned from a series of interesting but unconnected screens into bona fide research.

These wet-lab and computational skills have equipped me to take the lead on a research project seeking to answer the question: how can we assess the functional consequences of individual genetic variants at a large scale? This problem has clinical importance: patients with a "variant of uncertain significance" in a clinically actionable gene are often left in limbo, because low-throughput assays cannot match the pace at which genetic variants are uncovered.

High-throughput functional assays have the potential to scale to the challenge, but require methods to rapidly and precisely introduce genetic variants in relevant cell types. CRISPR base editors, which introduce point mutations at endogenous loci, seemed like a good candidate -- if they could perform well in pooled screens.

We first tested whether base editor screens could identify non-functional variants in a screening setting. My colleague and I conducted screens in five human cancer cell lines using a library that included all possible guide RNAs targeting a subset of essential genes. When I analyzed the data, I found that guides predicted to introduce splice site and nonsense mutations in essential genes were depleted relative to those predicted to introduce silent mutations. This result indicated that base editors were efficient enough to effectively identify nonfunctional variants in pooled, negative selection screens.

With the efficacy established, one clear use case was to identify non-functional variants in clinically relevant genes. I therefore conducted base editor screens tiling *BRCA1*, *BRCA2*, *TP53*, and other tumor suppressor genes for which non-functional variants are clinically detrimental. These screens identified missense and intronic mutations that scored strongly as non-functional, including both known pathogenic mutations (further validating the system's efficacy) and novel variants. The latter represent a "short list" worthy of follow-up; if these variants validate as non-functional in individual assays, they could eventually be re-classified as pathogenic, which would provide concrete guidance to patients with these variants.

Another idea was to use base editor screens to mutagenize drug targets and identify resistant mutations. I tested this approach by screening a library of all possible guide RNAs targeting an anti-apoptotic gene, *MCL1*, in the presence of an MCL1 inhibitor, such that drug resistant mutations would enrich. In order to visualize the hit mutations on the 3D structure of the MCL1 protein, I taught myself to use PyMOL, a Python-based molecular visualization program. Excitingly, I found resistant mutations within the small molecule binding pocket -- as we had hoped -- but also in regions of the protein not previously associated with drug resistance. Additionally, even though we were looking for resistance hits, we also found three guide RNAs, all within a few amino acids, that strongly sensitized cells to the inhibitor. The exact mechanism of these mutations remains unclear: although mutations in the binding pocket likely sterically hinder small molecule binding, how do mutations outside the binding pocket confer resistance? And what is the mechanism by which missense mutations in a drug target confer drug sensitivity? These questions may be beyond the scope of our current manuscript, but represent the type of problems that I would like to dig into during graduate school.

As I hope to pursue a career in research, my goals for graduate school are twofold: to obtain a strong training in the practice of science, and to develop expertise on a particular biological question. I believe MIT Biology's exciting faculty and strong focus on graduate education align with both of these goals. In particular, the first-year program sets MIT apart: the immediacies of experiments can obscure the big picture, and a semester of coursework would allow me to think broadly about the biological questions I'd like to explore. After that semester, I would like to rotate with Chris Burge, Rebecca Lamason, Stefani Spranger, and Sebastian Lourido. Although

these faculty tackle diverse biological problems, I am interested in the process of using large-scale techniques to perturb and dissect complex biological systems.

Entering the Broad, I sought to forge my own path through uncharted territory. Seeing this base editor screening project from start to (almost) finish has provided plenty of those exciting moments: alone in the tissue culture room, sensing the inklings of a fresh result for the very first time. And, as the project advanced, I also advanced my own skills: I learned how to code, how to contribute to a multi-group collaboration, and how to share our results with a large audience, first within the Broad and most recently at a conference at Cold Spring Harbor.

But perhaps most importantly, I've learned that some of the best moments are collaborative: sharing that fresh new result with a teammate, or watching our newest research associate skillfully explain the project to someone else, even though she's only been working on it for a month. And, although I took the lead on the project, the sheer number of people involved is staggering: the other research associates in our group, who conducted screens and validation experiments; the GPP production team, who produced liters (truly) of lentivirus; the computational biologist in our group, who wrote key scripts to annotate guides; our collaborators in David Liu's and JT Neal's groups, who developed and pressure-tested constructs; and my PI, who was absolutely invaluable in guiding every step of the project. So, yes, this project provided me with what I originally sought: the opportunity to advance my skills and the confidence that a career in science is right for me. But perhaps more importantly, my experience at the Broad demonstrated that collaboration is indispensable for science, a lesson that will shape my approach to research in graduate school and beyond.