

Goals for Fellowship Training and Career

My career goal is to be a leading scientist in either an academic research laboratory or comparable industry-based or non-profit research institution in the field of cancer biology. My long-term research interests involve the development of a comprehensive understanding of key genetic, epigenetic, and other regulatory mechanisms driving cellular identity and heterogeneity within cellular groups, tissues, and organs, particularly in the context of cancer and how this heterogeneity shapes tumor progression, therapeutic resistance, and ultimately clinical outcome. I believe that in order to understand this heterogeneity, novel statistical methods with user-friendly computational software must be developed to enable researchers to harness the power of big high-throughput sequencing datasets currently being generated. My goal for the F31 is to gain the skills and knowledge necessary for a lifetime career in the development of statistical methods and computational software for the analysis of high-throughput sequencing cancer data to uncover the biological mechanisms driving cancer development and prognosis. My advisor, Dr. Peter Kharchenko, and I have developed a training plan to work toward the following goals:

Build a strong quantitative and computational skill set: I have taken undergraduate and graduate courses that provide me with the necessary technical background for developing statistical methods and computational software including applied mathematics courses such as stochastic processes and Bayesian inference along with computational courses in data structures and machine learning. Throughout this fellowship, I will continue to learn statistical methods and computational approaches through hands-on research experience, attending conferences and meetings, and reading scientific literature. Specifically through this proposed research, I will learn to formulate problems analytically into a Bayesian framework, build Bayesian models, optimize model fitting, and assess model performance.

Strengthen background in the life sciences: In order to better interpret the biological significance of my computational analyses as well as improve communication with wet lab collaborators, I strive to become more familiar with topics in life sciences disciplines, with particular emphasis on immunology and cancer biology. I have taken undergraduate and graduate courses that have provided me with a strong biological background in molecular biology and genetics. In addition, I will attend the Immunology Seminar Series at Harvard Medical School and continue collaborating closely with Dr. Lili Wang of the Wu lab to strengthen my background in immunology and cancer biology. Dr. Wang has extensive experience as a trained immunologist working with chronic lymphocytic leukemia and her mentorship will provide me with a better understanding of the biological background and interpretation of my computational analyses.

Apply computational skills to gain biological insight from high-throughput sequencing data: I will apply the skills and knowledge acquired from the previous two goals to this proposed research project. This proposed research will also expose me to various sources of technical errors and considerations common to sequencing data that I might encounter again in the future in order to design and test new analysis pipelines as technological advances continue to develop.

Learn to independently lead and manage a scientific project: While my sponsors, collaborators, and fellow lab members will provide advice and suggestions as I complete the aims outlined in this project proposal, I will be in charge of developing and completing the proposed research. The experience I gain from managing this long-term project will be very valuable for my development as an independent researcher.

Manage and mentor students: Throughout my training, I will have the opportunity to mentor summer interns and more junior graduate students. I hope to gain experience managing and mentoring students by providing them with small projects related to the research outlined in this proposal such as applying the developed methods to novel datasets or creating simulations. Through mentoring students, I will develop skills needed to help train the next generation of scientists.

Improve communication and presentation skills: In order to effectively communicate my research with the broader scientific community, I strive to improve my scientific writing and oral presentation skills. In terms of scientific writing, I will continue gaining experience in preparing and submitting manuscripts to peer-reviewed journals. I will be practicing my oral presentation skills by presenting my research in regular lab meetings, program and institutional retreats, and local and national conferences. Furthermore, I will seek the mentorship of Dr. Nils Gehlenborg on proper visualization of my data. Dr. Gehlenborg has published extensively on data visualization and his mentorship will help me improve my graphical data representation and presentation skills.