Jeffrey P. Spence
Department of Genetics
Stanford University
jspence@stanford.edu
262-993-2435

October 25, 2023

Faculty Search Committee
Institute for Human Genetics
University of California, San Francisco
San Francisco, CA 94143

Dear Search Committee,

I am writing to apply for an Assistant Professor position in the Institute for Human Genetics at the University of California, San Francisco. I did my Ph.D. work in computational biology with Dr. Yun Song at the University of California, Berkeley and am currently a postdoctoral researcher in Dr. Jonathan Pritchard's group at Stanford University. My research focuses on understanding the genetics of complex traits and human disease by combining and developing techniques across statistical genetics, population genetics, and machine learning.

One of the major themes of my research is increasing the biological complexity of the models used in statistical genetics and population genetics, while developing statistical inference methods that can scale to modern, massive datasets. I have taken this approach in several projects, always centered on understanding the causes and consequence of human genetic variation. This has ranged from population genetics to statistical genetics. In population genetics, I developed methods to infer the demographic history of a population from genetic variation data and methods to account for demographic history when estimating rates of recombination across the genome. In statistical genetics, I have built more realistic models of how variants affect disease traits and have used these models to build improved polygenic scores for disease prediction. I have often needed to develop new statistical techniques in order to push models to be more biologically realistic while keeping inference scalable, resulting in methods that are of independent interest to the machine learning community.

My future research will focus on understanding disease by building complex, realistic, evolution-informed models and fitting those models to biobank-scale data using advances from machine learning. Statistical genetics traditionally approaches problems by treating observed genetic variation as fixed and then testing for associations between genetic variants and disease. That is, statistical genetics typically ignores what causes observed genetic variation. On the other hand, the cause of observed genetic variation is exactly the study of population genetics. Ultimately, statistical genetics is leaving information on the table — genetic variation *is* informative about disease: natural selection acts on disease-associated variants shaping genetic variation. By combining ideas from these complementary fields, I will deepen our understanding of human health.

One major challenge for combining models from population genetics and statistical genetics is that such models quickly become analytically intractable. Previously, I worked to push the biological complexity of models while keeping inference tractable by using pen-and-paper analytical techniques. This often took

the form of deriving bespoke inference schemes or working to exploit structure specific to each particular model. In ongoing work, I have shown that using supervised machine learning on data simulated from these types of models has a solid statistical underpinning. This opens up a whole world of possibilities by allowing us to perform valid statistical inference on incredibly complex models by "black boxing" the mathematical analysis that would traditionally need to be done using pen and paper. My research group will focus on extending these ideas and applying this framework to models combining statistical genetics and population genetics approaches to further our understanding of disease genetics.

I am particularly excited by the environment at UCSF in terms of the incredible diversity of world class research, from human genetics, to cutting-edge technology development, to having a translational impact on human health in the clinic. I thrive on collaboration as evidenced by my contributions to publications spanning diverse areas from ecology to human history to human health. In these collaborations, I provide expertise in building and performing inference on complex, but interpretable statistical models, as well as the population genetics of complex traits and statistical genetics methodology. I would be extremely interested in collaborating with the labs in the IHG as well as across UCSF.

On a personal note, I have lived in the Bay Area for the past ten years, including several years in Mission Bay, just a few blocks from the UCSF Mission Bay campus. I cannot imagine living or working anywhere else in the world, and turned down faculty offers from Columbia, Penn, and OHSU last cycle in order to stay in the Bay Area.

Thank you for considering my application, and please feel free to contact me with any additional questions.

Sincerely,

Jeffrey P. Spence