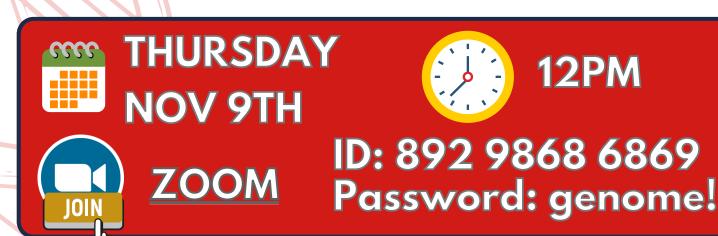


Genetics, Bioinformatics, & Systems Biology Colloquium

presents

Sara Mostafavi, PhD

Associate Professor
University of Washington





SEQUENCE-BASED DEEP LEARNING MODELS FOR UNDERSTANDING GENE REGULATION AND DISEASE GENETICS

The mammalian genome contains several million cis-regulatory elements, whose differential activity marked by open chromatin determines cellular differentiation. While the growing availability of functional genomics assays allows us to systematically identify cis-regulatory elements across varied cell types, how the DNA sequence of cis-regulatory elements is decoded and orchestrated on the genome scale to determine cellular differentiation is beyond our grasp. In this talk, I'll present our work on applying and interpreting sequence-based deep learning models to derive an understanding of the relationship between regulatory sequence and cellular function in the context of immune cell differentiation. I will then describe how these models can be applied to understand the impact of unseen genetic variation across diverse cellular phenotypes, while discussing challenges when models are applied in such out-of-sample prediction tasks. Finally, I will discuss our recent work that improves the ability of sequence-based-models to make predictions for unseen genetic variations. In summary, our work shows that sequence-based deep learning approaches can uncover patterns of immune transcriptional regulators that are encoded in the DNA sequence, and can provide a powerful in-silico framework to mechanistically probe the relationship between regulatory sequence and its function.



