



Department of Biomedical Informatics Special Seminar



Zijun (Frank) Zhang, PhD
Assistant Professor, Division of AI in Medicine, and
Department of Computational Biomedicine at
Cedars - Sinai, Los Angeles, CA

Genomic AI: Bridging Foundation and Specialist Models

Wednesday, June 26, 2024

2:00pm - 3:00pm

Join Zoom Meeting:

<https://stonybrook.zoom.us/j/95186108918?pwd=gplhxRvYRfALEHSv07LQYaQ8MvraG1.1>

Meeting ID: 951 8610 8918 Passcode: 206534

Abstract: The advancement of large language models (LLMs), known as foundation models, provides exciting opportunities for understanding the regulatory grammar of the genome. Despite their success in natural and protein language processing, the application of LLMs to genomics remains challenging. In this talk, I will discuss a novel strategy of genomic artificial intelligence (AI), by bridging foundation and specialist genomic models. Focusing on using AI to understand the personalized genetic variations, I will explore three projects of developing new AI methods. First, I will present how we fine-tuned protein and DNA LLMs by leveraging pretrained genomic knowledge in these foundation models for disease-specific rare variants, with successful replication of our model performance in ClinVar. Second, I will explain how the performance of specialist models, in contrast to foundation models, is dependent on the design of neural architectures; and how we employ a training-free model design method to automatically search for accurate deep models across various data modalities in genomics and biomedicine. Finally, I will discuss our latest efforts in using specialist models as encoders to align DNA regulatory grammars for natural language grammars within foundation models. By integrating foundation and specialist models, we model gene expression across 1 million DNA bases by Llama3, demonstrating a new paradigm of domain-shift computing in genomic AI.

Bio: Dr. Zhang is an assistant professor in the division of AI in medicine and the department of computational biomedicine at Cedars-Sinai. He has longstanding interests and expertise in developing deep learning methods for genomics, multiomics and biomedicine. His work on powerful deep learning-based computational methods has been published in renowned journals, such as Nature Methods, Nature Machine intelligence, and Nature Computational science. He is a selected nominee for the Blavatnik Award for young scientists in 2022 and the recipient of the Winnick Award for clinical investigators in 2023