

# CARNEGIE MELLON UNIVERSITY

## BME 2025 SPRING SEMINAR SERIES



### Engineering cell state-specific gene regulation with AI and synthetic biology



#### PRESENTED BY

**Sebastian Castillo Hair, Ph.D.**

Postdoctoral Researcher  
Department of Biomedical  
Engineering  
University of Washington

#### SCHEDULE

**Doherty Hall (DH) 2315**

**Tuesday,  
February 25, 2025  
(11:00-12:00 PM)**

Cells across tissues, developmental stages, and disease conditions adopt distinct intracellular states – epigenomic, transcriptomic, and proteomic profiles – to compartmentalize function in time and space. The ability to write DNA- and RNA-encoded programs that sense and interface with cellular states has transformative potential for biotechnology, for example in designing gene therapies with cell type- and disease-specificity to minimize off-target effects. However, limited understanding of how cellular processes respond to cellular states remains a major bottleneck. Machine learning (ML) models that learn the regulatory grammar of cell state regulation from large datasets offer a powerful tool to overcome these barriers. In this talk I will present advances in combining ML and high throughput experiments to study and engineer DNA and RNA sequences that recognize cell types, tissues, and developmental stages. First, I will focus on enhancers – DNA sequences that regulate transcription with cell type-specificity. We trained ML models on databases of genomic accessibility – an imperfect proxy of enhancer activity that is available for hundreds of cell types and tissues – and used them to design enhancers targeted to multiple cell types. Experimentally, these enhancers drove stronger and more specific gene expression than endogenous controls when tested on 10 human cell lines, including from muscle, liver, breast, and embryonal tissues, and even in vivo in mouse retinas. In the second part, I will focus on mRNA sequence engineering. We used ML predictors of translation to optimize performance of an mRNA-delivered gene editor, and developed predictors of mRNA degradation in response to microRNAs—non-coding RNAs that are differentially expressed across cell types. Using these, we designed mRNAs with programmable specificity across 10 cell lines. Finally, I will focus on gene regulation during the dramatic cell state transitions that occur during development. By training ML predictors of chromatin accessibility and mRNA regulation during zebrafish embryogenesis and applying explainable AI methods, we uncovered regulatory grammar driving differentiation and gene expression dynamics. Our work shows the potential of combining ML, high throughput experiments, and synthetic biology to decode and engineer cell state-responsive biological systems.