Name: Chaitanya Srinivasan
PI Name: Andreas Pfenning

Email: csriniv1@andrew.cmu.edu
PI Email: apfenning@cmu.edu

Machine Learning Prioritization of Cell Type-Specific Enhancers in the Nucleus Accumbens Implicated in Addiction Predispostion

Chaitanya Srinivasan[1,2], Madeline K. Fish[4,5], BaDoi N. Phan[1,2,3], Chen Fu[4,5], Snehal Sambare[4,5], Marianne L. Seney[6], Ryan W. Logan[4,5], Andreas R. Pfenning[1,2]

[1]Ray and Stephanie Lane Computational Biology Department, Carnegie Mellon University; [2]Neuroscience Institute, Carnegie Mellon University; [3]Medical Scientist Training Program, University of Pittsburgh School of Medicine; [4]Department of Psychiatry, University of Massachusetts Chan Medical School; [5]Department of Neurobiology, University of Massachusetts Chan Medical School; [6]Department of Psychiatry, University of Pittsburgh School of Medicine

Substance use disorders (SUDs) and overdose are increasingly a major public health problem, yet the development of effective therapeutic interventions remains challenging on multiple axes. From a genetic perspective, SUDs are polygenic and heritable, with underpinnings in common genetic variants that share liability with non-substance-related psychiatric traits. On a second axis, the pathophysiology of addiction-related behaviors can be modulated through epigenetic mechanisms in noncoding, cis-regulatory elements (CREs) of the genome. Genetic risk variants can disrupt transcription factor binding in CREs, resulting in altered downstream gene expression and neural circuitry. Finally, SUDs involve an interplay of several cell types and brain regions in different underlying addiction behaviors. While tremendous prior research has linked addiction behaviors to the NAc, an important hub of reward circuitry, their cell type-specific regulatory bases remain an open area of inquiry crucial to improving therapy development.

We measured high quality, same cell single nucleus RNA-seq and ATAC-seq from postmortem human NAc (N=24 subjects). Leveraging these data, we annotated cell types of the NAc and developed convolutional neural networks to predict cell type-specific CRE activity from DNA sequences underlying putative CREs. We applied these models to 1) prioritize NAc subtype-specific CREs, 2) target NAc subtypes using AAVs in wild type mouse, and 3) predict the functional impact of SUD risk variants in CREs. We interpret our models by identifying classes of transcription factor binding sites driving cell type-specificity in NAc subtypes. Our findings provide insight into regulatory mechanisms underlying SUDs and a platform for targeted gene therapy development.

DRAFT