Name: Andreas Pfenning Email: apfenning@cmu.edu

What can bats, capybaras, and tree shrews tell us about the genetic basis of addiction?

BaDoi Phan[1,2], Alyssa Lawler[3], Jing He[4], Irene Kaplow[1], Daniel Schaffer[1], Ashley R Brown[1], Zoonomia Consortium, William Stauffer[4], Andreas R Pfenning[1,3]

[1]Department of Computational Biology, School of Computer Science, Carnegie Mellon University; [2]Medical Scientist Training Program, School of Medicine, University of Pittsburgh; [3]Department of Biological Sciences, Mellon College of Sciences, Carnegie Mellon University; [4]Department of Neurobiology, School of Medicine, University of Pittsburgh

Hundreds of loci with thousands of genetic variants have been associated with addiction-related behaviors, which are largely thought to act at distal regulatory elements and in tissue- and cell type-specific contexts. For many human diseases, causal genetic variants tend to be conserved across mammalian evolution, providing a means to prioritize which candidate genetic variants within a given locus are likely to be driving disease predisposition. However, for genetic variants associated with addiction-related behaviors, we only find a modest enrichment of conserved nucleotides. To better prioritize and interpret addiction-associated genetic variation, we create a new conservation metric, Cell-TACIT, that combines cell type-specific epigenetics with machine learning. We matched 8 orthologous cell types in the caudate nucleus of the striatum across human, rhesus macaque, and mouse with open chromatin assays. We then trained convolutional neural network models to predict active open chromatin. We apply these models, the Cell Type-Activity Conservation Inference Toolkit (Cell-TACIT), to predict the open chromatin activity of orthologous conserved DNA of 240 mammals. Genetic variants associated with smoking session show an 18-fold enrichment in conserved nucleotides based on the traditional phyloP score (adjusted p=0.33), but a 109-fold enrichment where Cell-TACIT predicts the region is conserved open chromatin in D1 medium spiny neurons across mammals (adjusted p=0.01). Overall, we find that mapping the conserved cell-type regulatory patterns refines the list of candidate causal variants in regions associated with a variety of addiction-related traits beyond conservation or human cell type-specific open chromatin alone.