Submitter Name: Elissa J. Chesler
Submitted email: elissa.chesler@jax.org

Interpreting human genetics of SUD using multi-species heterogeneous functional genomics data

Elissa J. Chesler[1], Timothy Reynolds[1,2], Jason A. Bubier[1], Emma Johnson[3], Vivek Philip[1], Jake Emerson[1], Rohan Palmer[4], Arpana Agrawal[3] and Erich J. Baker[2]

[1]The Jackson Laboratory; [2]Computer Science Department, Baylor University; [3]Department of Psychiatry, Washington University, St. Louis; [4]Psychology Department, Emory University

Human genetic studies of substance use disorders (SUD) have begun to identify variants associated with various characteristics and categories of SUD. Prioritizing these variants and identifying their mechanistic roles and possible therapeutic potential in SUD requires substantial additional information. Model organism studies provide extensive neurobiological depth and genomic breadth but need to be integrated with human genetic studies in a meaningful way to relate variants, genes, and gene products to precise mechanisms of addiction. The GeneWeaver system provides a data model for multi-species functional genomic data integration that is implemented for access by a suite of modular APIs for highly customizable and reproducible user initiated analyses. The model has recently been expanded for variant specific analyses. Using this data resource and analysis tool suite, we demonstrate three applications, 1) the prioritization and characterization of GWAS variants based on model organism evidences, 2) the identification of mouse behavioral assays that best resemble a human psychiatric disorder through Network Enhanced Similarity Search, 3) the use of model organism data to improve the predictive power of polygenic risk scores for SUD. This extensible software system can support integration and interoperation of addiction genetics and genomics summary data from heterogeneous sources for these and other applications.