Identifying Addiction-Associated Genes: Analyzing RNA-Seq Data from Brain Tissues to Reveal Genetic Drivers of Addiction Phenotypes

May 5, 2023

Authors: Anna Saboe, Ryan Mower, Muhammad Abuzar, Grace Walker, Greg Shobert, Luke Bessant

University of Minnesota
Driven to Discover℠

Background

Large gap in knowledge regarding genetic basis of addiction

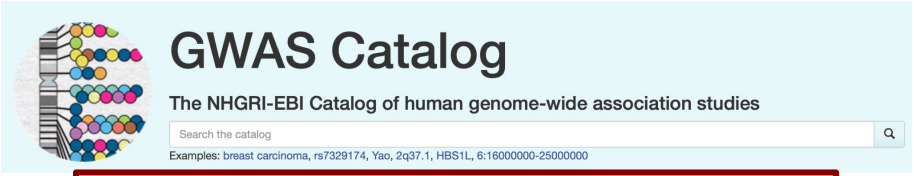
Initial interest in understanding the relationship between age and addiction phenotypes

Can genetics predict the risk of addiction? Particularly in regions of the brain important for addiction: basal ganglia, amygdala, prefrontal cortex

Are certain genes associated with addiction expressed differently across areas of the brain impacted by addiction related disorders?

Methods

Obtaining Data:

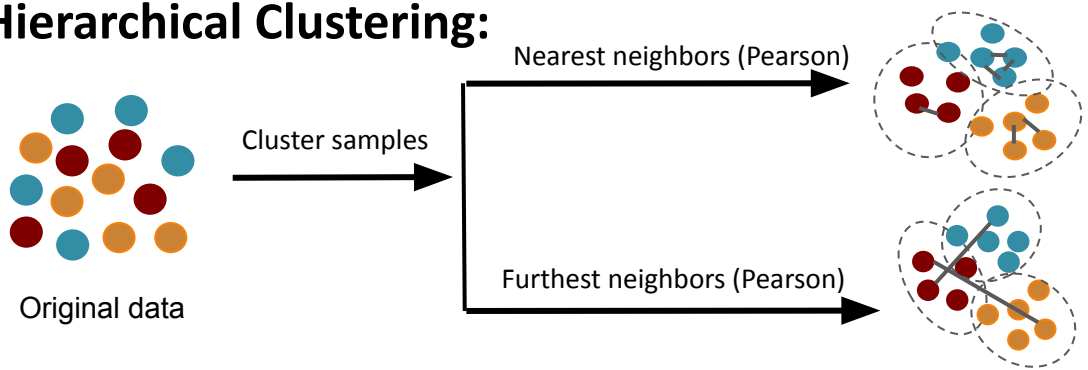


SNPs mapped to genes correlated with drug dependence

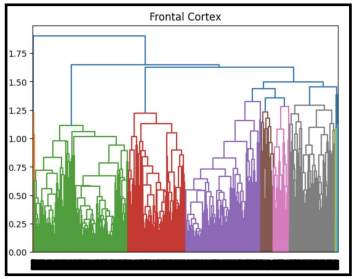
GTEx Portal

TPM values from RNA-seq data from basal ganglia (n = 205), prefrontal cortex (n = 209) and amygdala (n = 152) tissues

Hierarchical Clustering:



Original data



Generate dendrograms

Methods

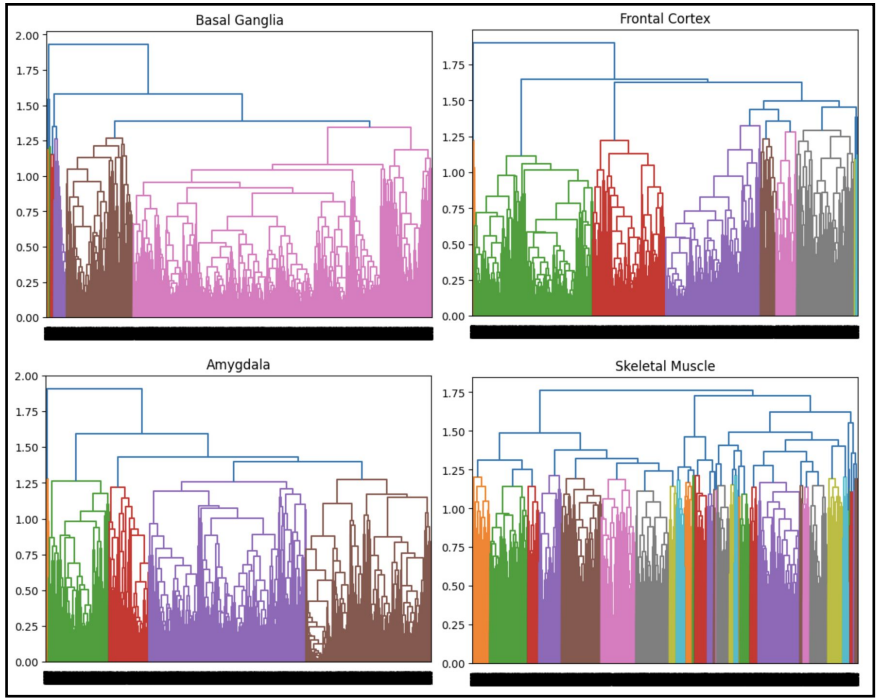
Clustering Analysis:

1. Select number of clusters that will result in a cluster mean of 20-100 genes
2. Determine cluster size distribution
3. Search through clusters to determine which contains highest number of genes associated with alcoholism

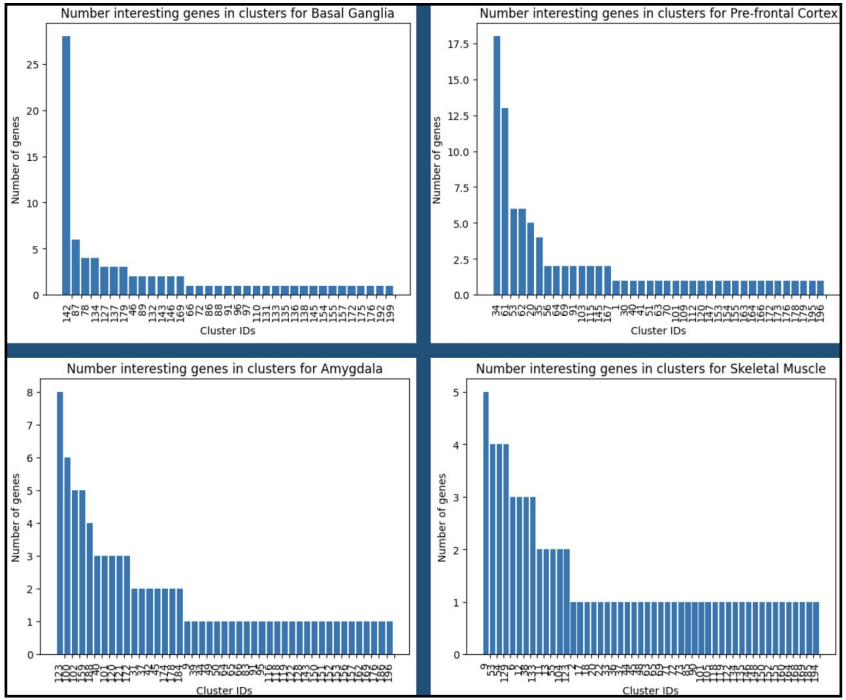
Gene Ontology (GO) Enrichment:

1. PANTHER tool
   - ➔ Input list of genes from most interesting cluster of each tissue type
   - ➔ Compare to *Homo sapiens* reference gene list
2. Test for biological function, molecular component and cellular processes
   - ➔ **Fisher's Exact Test** : used to examine the significance of the association between reference gene set and our cluster set
   - ➔ **False discovery rate (FDR)** : used to determine significance of features and reduce likelihood of false positives

Results: Hierarchical Clustering



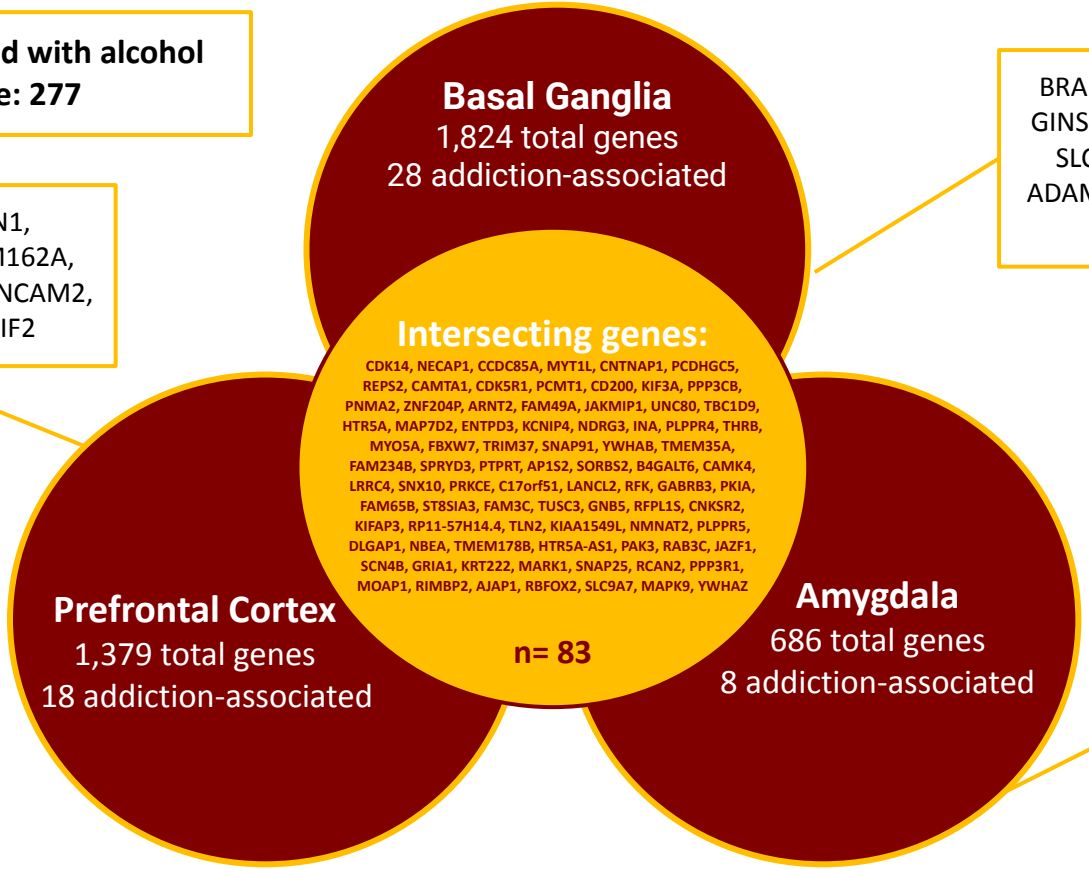
Dendrograms generated from clustering



Histograms displaying gene count distribution across clusters

Results

Total SNPs associated with alcohol dependence: 277



BRAP, SPATS2L, FTO, OTUD1, NCALD, GINS3, PCNX1, WDR7, NR4A2, MTIF2, SLC6A15, KIAA0040, PPFIA2, RB1, ADAM17, MARVELD1, HMG20A, FAF1, PCDH17, BBS4, NFIA

PPFIA2, MBNL2, NRXN1, TMEM14A, FAM81A, FAM162A, ANKS1B, CLCN4, CAMTA1, NCAM2, HMG20A, MGAT5, MTIF2

PPFIA2, MTIF2, MICU3, BRAP, FAM81A, LRTM2, CLCN4, FSTL5, CAMTA1, OPCML, NCAM2, FAF1, PKNOX2, SLC6A15, ZNF532

Results: GO Enrichment

Tissue	Biological Process	Molecular Function	Cellular Component
Basal ganglia	Synaptic vesicle budding	Phosphatidylinositol-3-phosphate phosphatase activity	Dendritic spine head
	FDR: 2.21E-02	*FDR: 4.87E-02*	FDR: *3.30E-02*
Prefrontal cortex	Neurotransmitter receptor localization to postsynaptic specialization membrane	Voltage-gated monatomic ion channel activity involved in regulation of presynaptic membrane potential	Postsynaptic intermediate filament cytoskeleton
	FDR: *1.55E-03*	FDR: *5.76E-03*	FDR: *1.24E-02*
Amygdala	Postsynaptic density structure maintenance	Synaptic receptor adaptor activity	Primary dendrite
	FDR: *3.55E-02*	FDR: *2.07E-02*	FDR: *3.53E-02*
Genes shared across all 3 tissues	Synapse organization	*No significant enrichment*	Kinesin II complex
	FDR: *3.43E-02*	FDR: *N/A*	FDR: *1.94E-02*

Conclusions

Clusters containing the highest number of addiction-related genes showed enrichment for biological elements relating to neural signaling

- Synapses → central to reward processing and plasticity, commonly impaired by addiction and relapse *(Kauer & Malenka, 2007)*
- Dendrites → vital contributions to neuron physiology and synaptic plasticity *(Spiga et al., 2014)*
- Kinase activity → addiction alters common network of signaling pathways relating to specific protein kinase subsets *(Lee & Messing, 2008)*
- Kinesin → postulated to be a transcriptional target for the regulation of of drug-induced changes to synaptic plasticity

Genes within our significant clusters *not previously linked to addiction* could provide novel avenues for further research into how all above processes are influenced by addiction

Future Directions

1

Broaden our dataset

- Greater representation of younger populations
- Other gene expression measurements
- Increase number of samples

2

Investigate other addiction phenotypes

- Nicotine
- Cannabis
- Opioids
- Codependence (addiction to multiple substances)

3

Further optimize clustering methodology

- Average linkage
- K-medoids

4

Investigate impact of different distance metrics on the dataset

- Rank Magnitude
- Euclidean distance

References

Kauer, J. A., & Malenka, R. C. (2007). Synaptic plasticity and addiction. Nature Reviews Neuroscience, 8(11), Article 11. https://doi.org/10.1038/nrn2234

Lee, A. M., & Messing, R. O. (2008). Protein kinases and addiction. Annals of the New York Academy of Sciences, 1141, 22–57. https://doi.org/10.1196/annals.1441.022

Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. Jan 2019;47(D1):D419-D426

Solomon BD, Nguyen AD, Bear KA, Wolfsberg TG. Clinical Genomic Database. Proc Natl Acad Sci U S A. 2013 May 21. [Epub ahead of print] [PubMed]

Spiga, S., Mulas, G., Piras, F., & Diana, M. (2014). The “addicted” spine. Frontiers in Neuroanatomy, 8, 110. https://doi.org/10.3389/fnana.2014.00110

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this report were obtained from the GTEx Portal on 04/01/23.