How Data Science is Helping Us Understand the Brain in HIV and Opioid Use

COGS 108 Spring 2025
Jason Chen
Week 10

xic007@ucsd.edu

OH: Thu, 3-5 PM

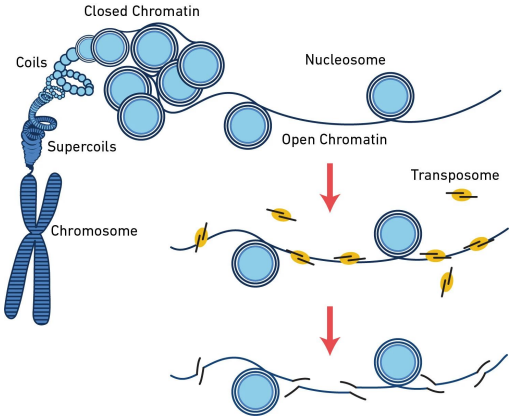
Due dates

- Final project due next Wednesday, Jun 11 @ 11:59 PM

SCORCH Objectives

SCORCH
Single Cell Opioid Responses in the Context of HIV

- Create single-nucleus transcriptome databases to identify rare cell types and key cellular populations, advancing understanding of CNS effects related to SUD and HIV.



Why Study the Brain? Why HIV and Opioids?

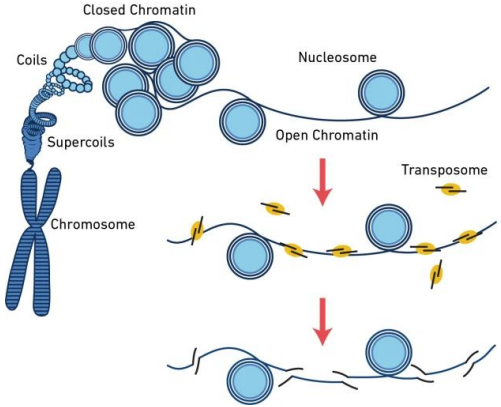
- The brain is *super complex* — and it's impacted by diseases in subtle, surprising ways.
- HIV affects the immune system, but also causes brain issues like memory loss or depression.
- Opioids can also rewire how brain cells communicate, especially those tied to emotion.
- Studying both together helps us understand how they might *interact* in the brain.

Fact: Over 39 million people live with HIV worldwide, and millions more are affected by the opioid crisis.

The Data We Collect

- Single-Nucleus Multiomic Data
  - RNA-seq: What genes are active in a cell
  - ATAC-seq: What regions of DNA are "open" and possibly active

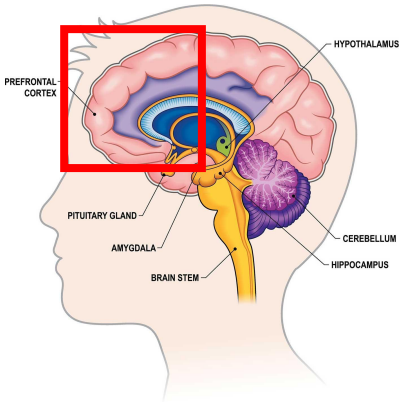
- Why they matter:
  - They give us a full picture of both gene expression and regulation



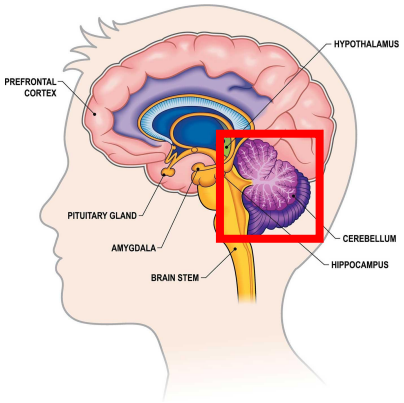
* Work in progress, not published

UCSD SCORCH

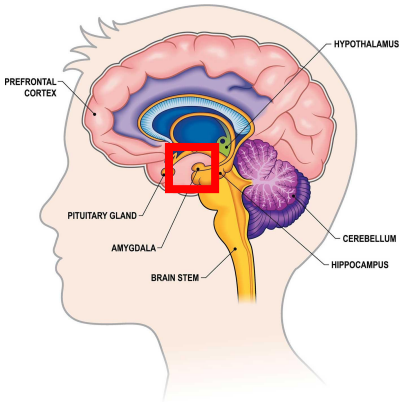
Three Brain Regions



PFC



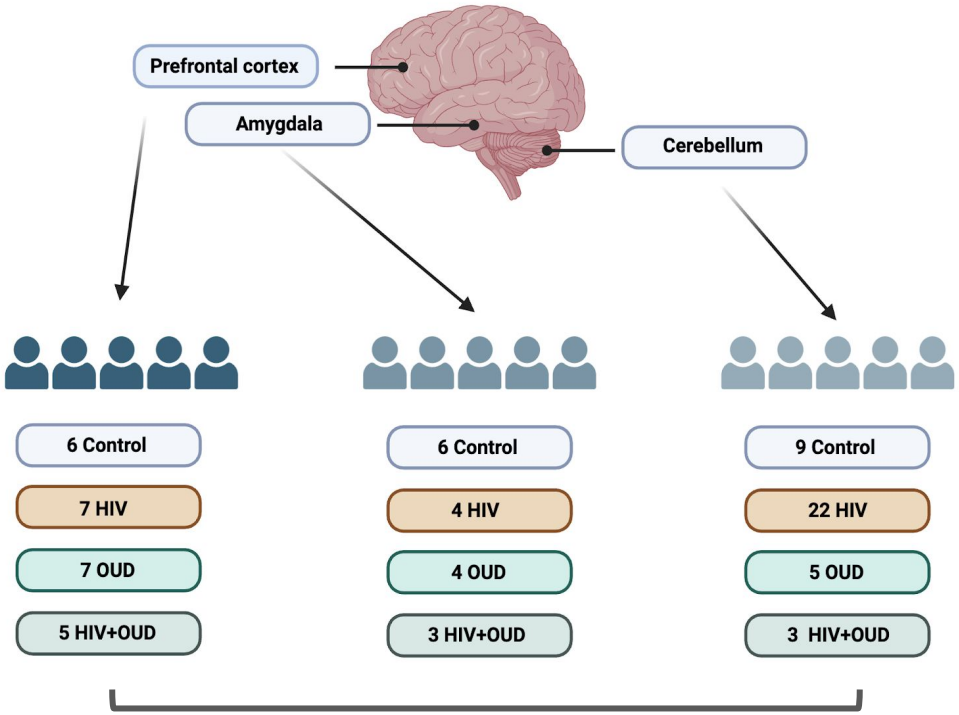
Cerebellum



Amygdala

* Work in progress, not published

Dataset Overview: 48 donors across 3 brain regions



470,000 high-quality multiome (RNA+ATAC) profiles

Dr. Tariq Rana

Dr. Allen Wang

Dr. Kyle Gaulton

Justin Buchanan

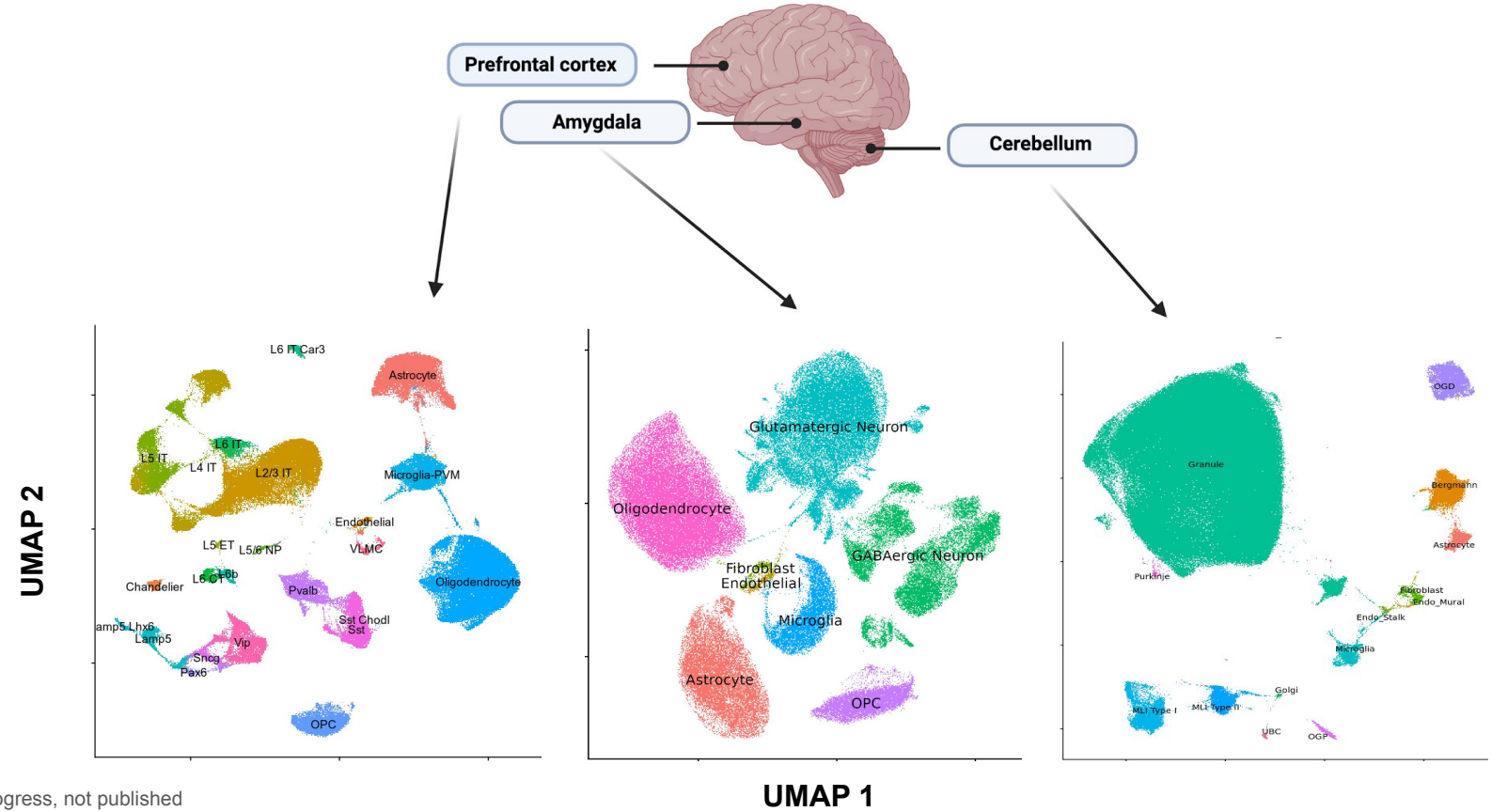
Abbey Green

Jason Chen

Tanmayi Vashist

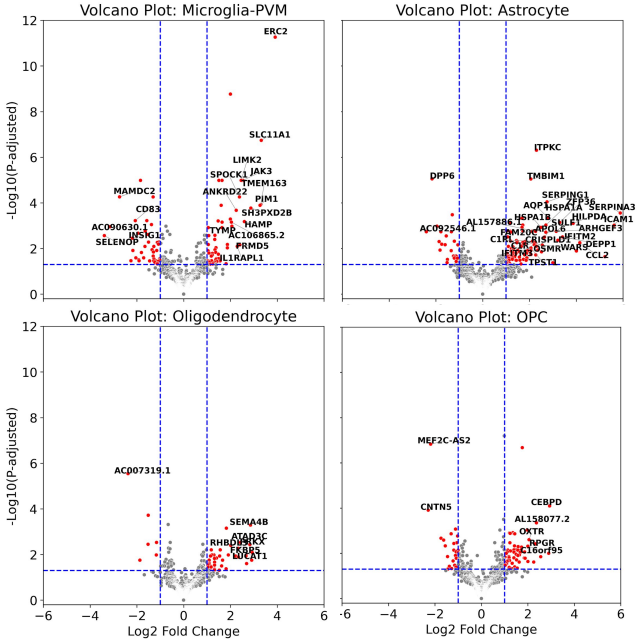
* Work in progress, not published

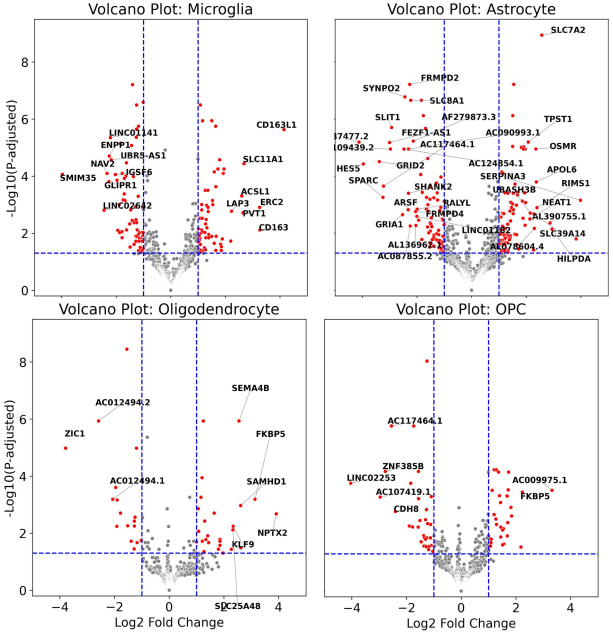
Dataset Overview: 45 donors across 3 brain regions

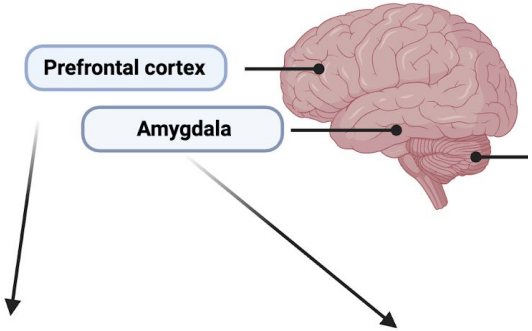


* Work in progress, not published

HIV vs Control DESeq2 results



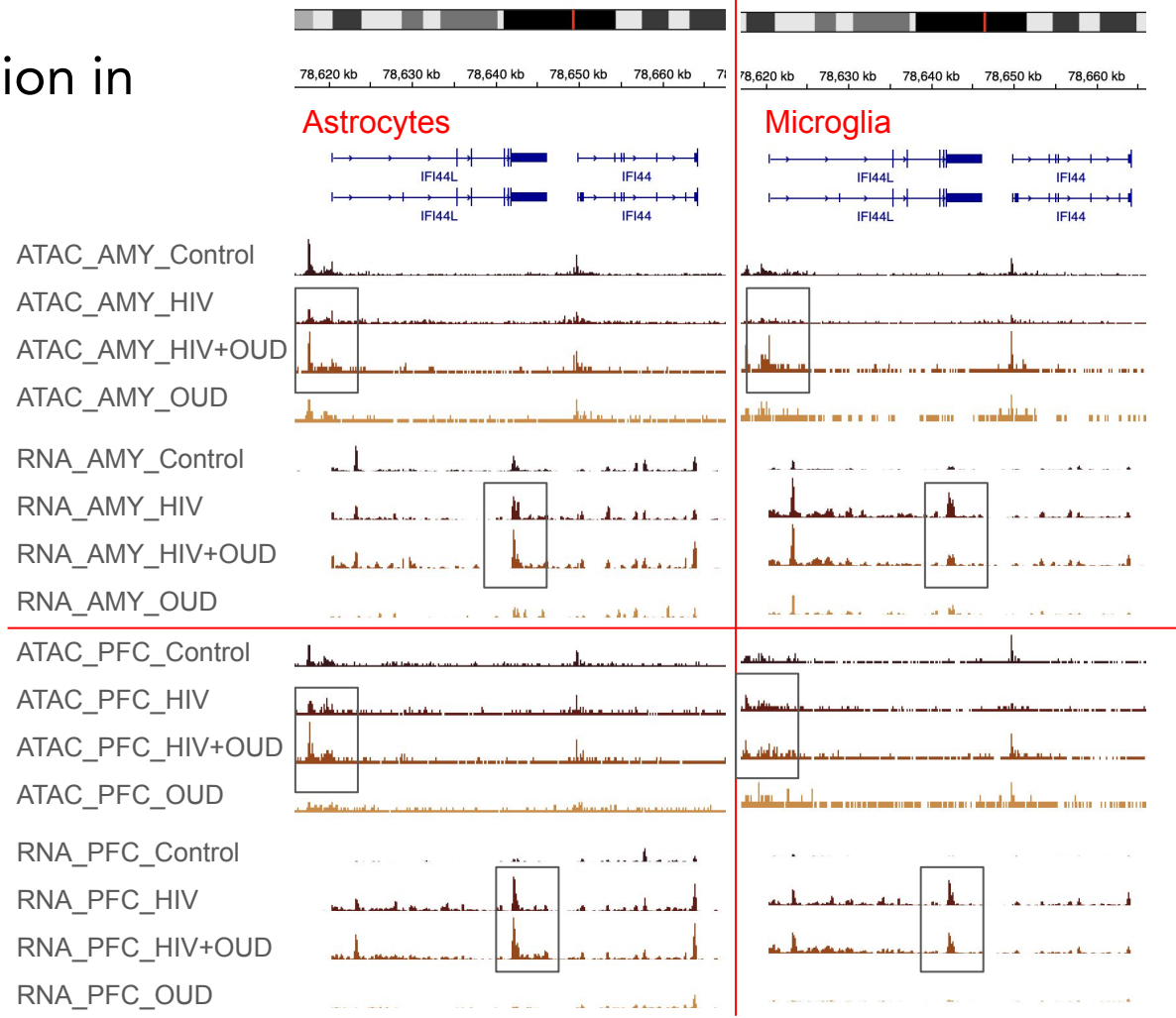




* Work in progress, not published

IFI44L shows higher expression in HIV+ samples

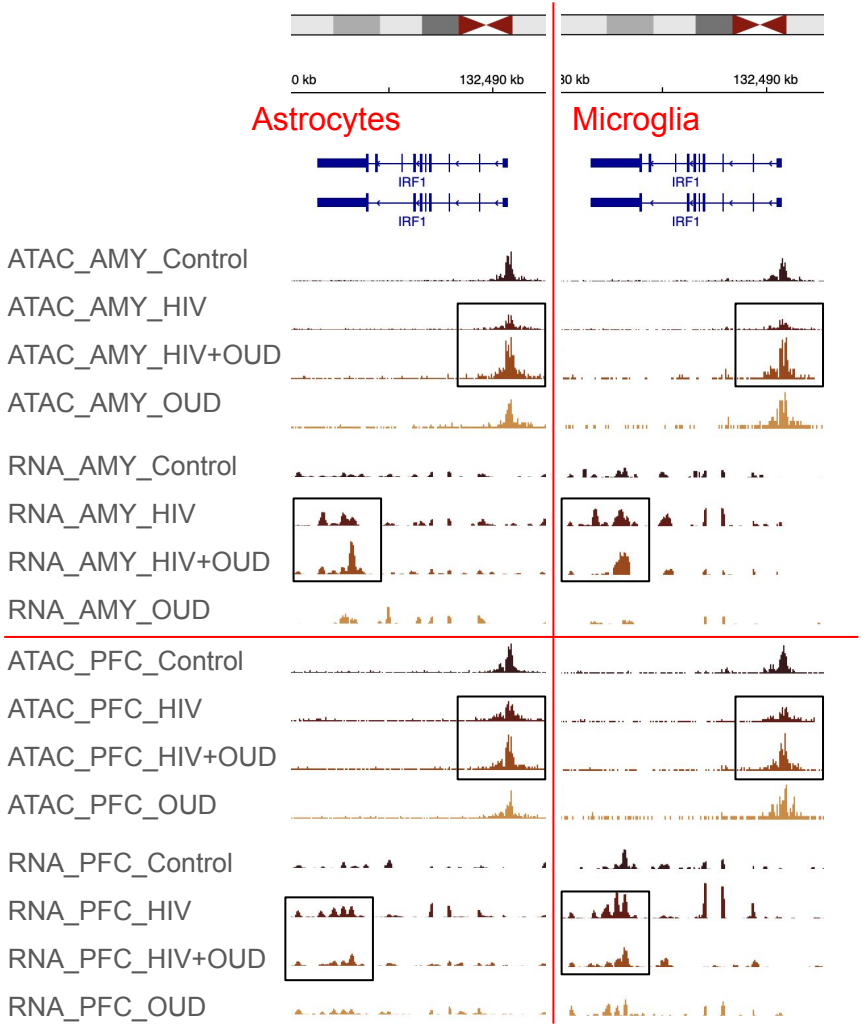
PFC Microglia: Wilcoxon test indicates significant difference in HIV (*FDR <.05*) and HIV+OUD (*FDR <.05*) IFI44L gene activity scores compared to Controls



* Work in progress, not published

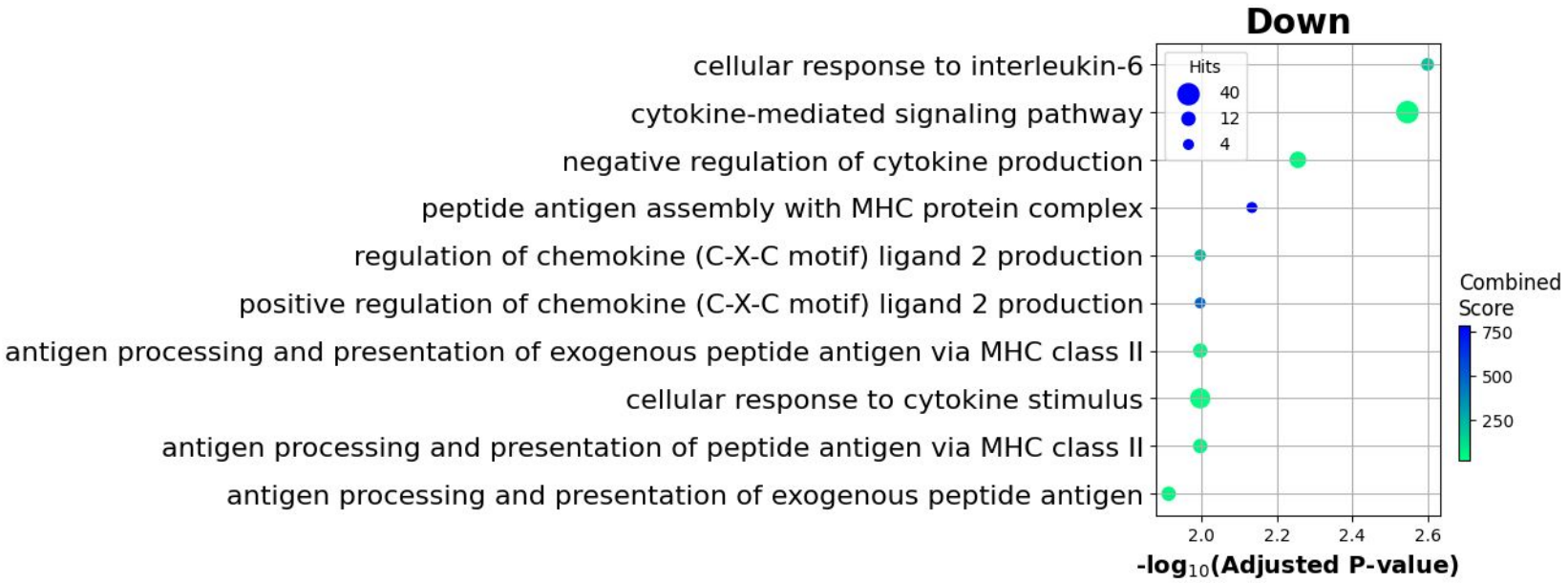
IRF1 shows higher expression in HIV+ samples

PFC Microglia: Differentially accessible peak (FDR = .0004) at the end of IRF1, HIV vs Control



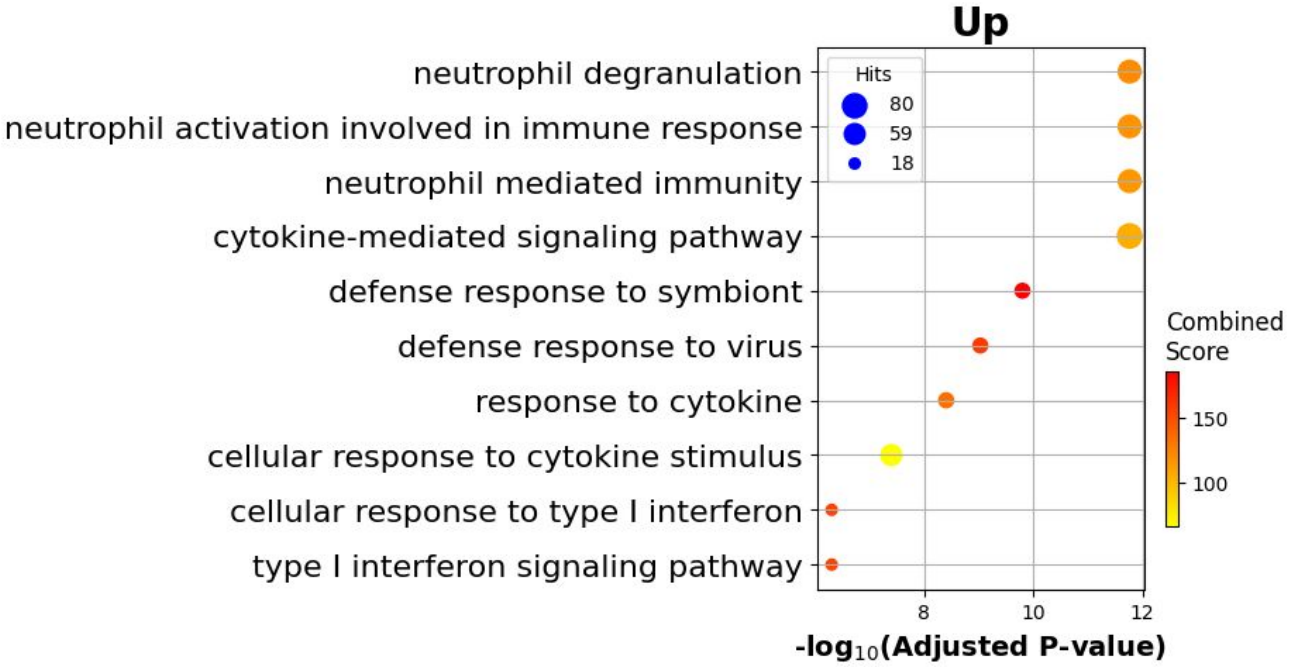
* Work in progress, not published

GSEA for Microglia Suggests suppressed antigen presentation and cytokine regulation



* Work in progress, not published

GSEA for Microglia Suggests compensatory immune activation

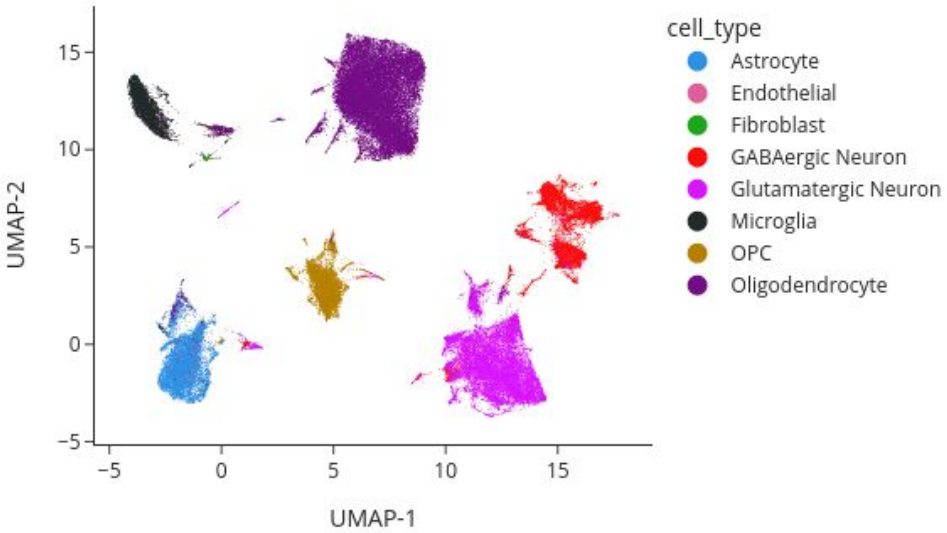


* Work in progress, not published

ATAC Data Preprocessing

- Used SnapATAC2 to load each ATAC fragment file into AnnData objects
- Used generated AnnData objects to create merged AnnDataSet
- Performed QC & Filtering
  - Kept cells with > 1000 fragments and > 5 TSSE
- Metadata transfer & filtering based on RNA data

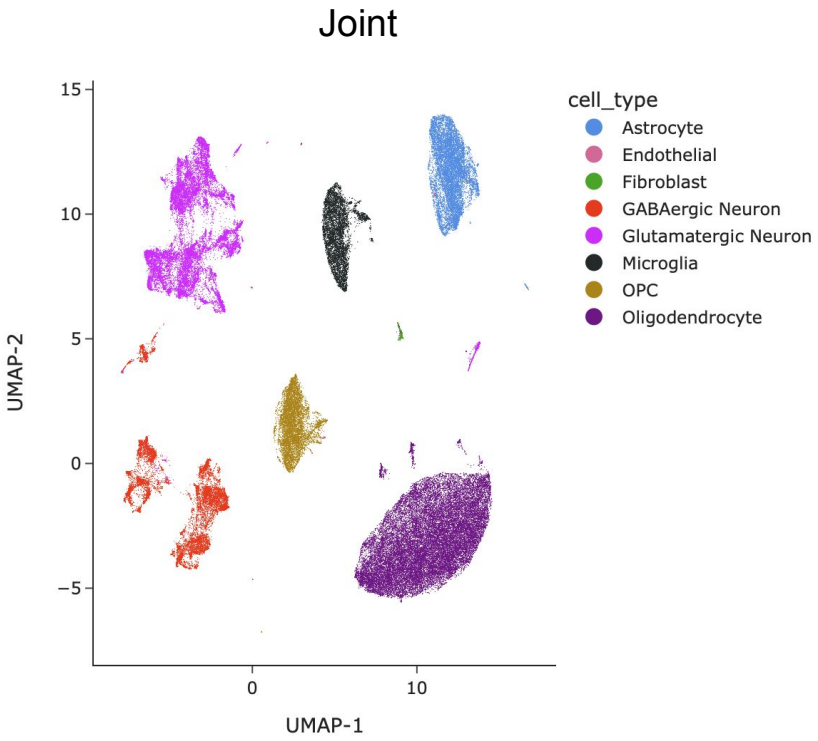
UMAP of ATAC



* Work in progress, not published

ATAC Data Preprocessing

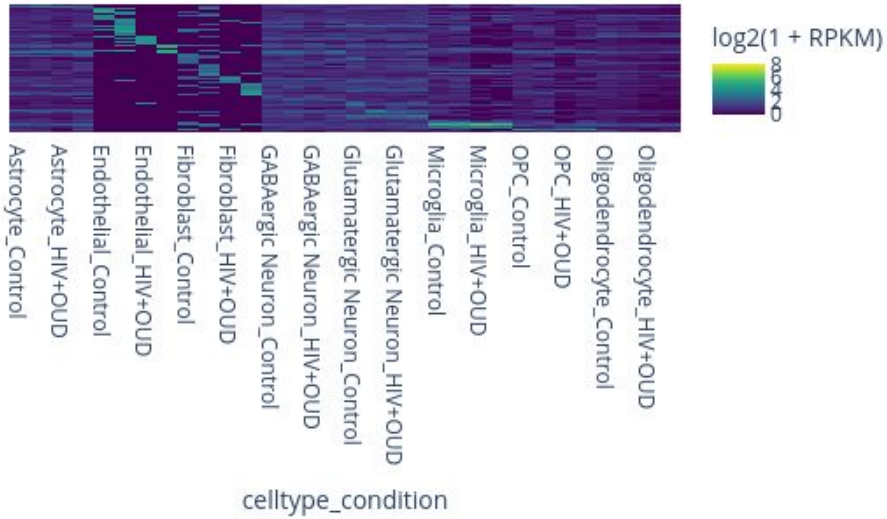
- Integrated RNA data to generate joint UMAP
- Batch correction with harmony



* Work in progress, not published

Peak Calling

- Grouped cells according to the “celltype_condition” label and then call peaks with MACS3 in each group
- Merged peaks
  - Kept most significant peaks for each overlapping peak set

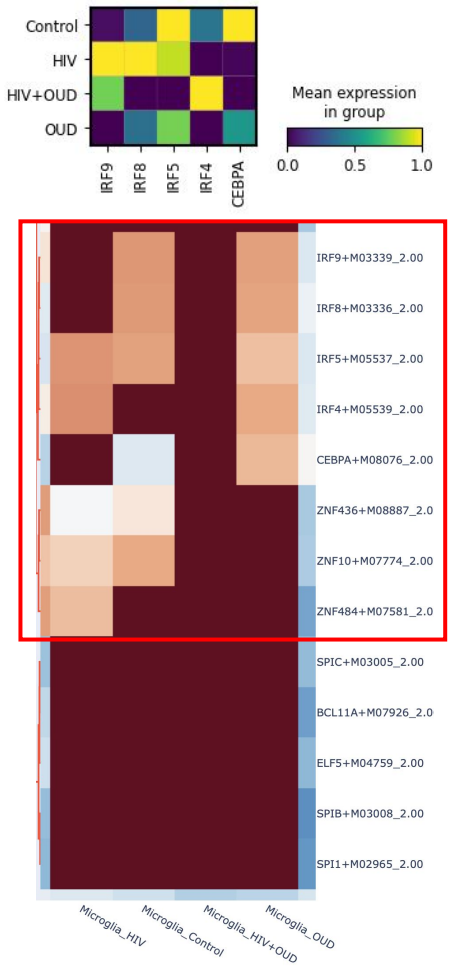


Heatmap of Marker Peaks

* Work in progress, not published

Motif Enrichment

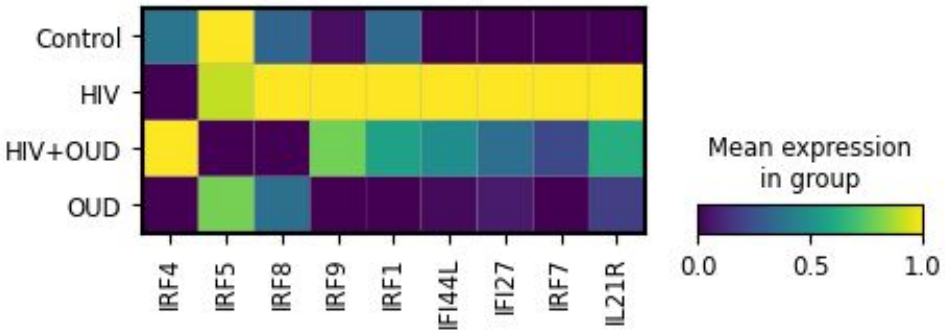
- IRF4, IRF5, IRF9 (Interferon Regulatory Factors):
  - major regulators of immune response and inflammation
  - suggests a strong antiviral microglial activation, likely driven by the interferon signaling pathway in response to HIV infection (also suggested by GSEA)
- CEBPA (CCAAT Enhancer Binding Protein Alpha):
  - a regulator can control the development and differentiation of microglia, also associated with inflammation (Zhou et.al, 2019)
  - suggests that HIV infection skews microglial differentiation, potentially toward a pro-inflammatory state



* Work in progress, not published

Interferon Regulatory Factors (IRFs) are Differentially Regulated in Microglia

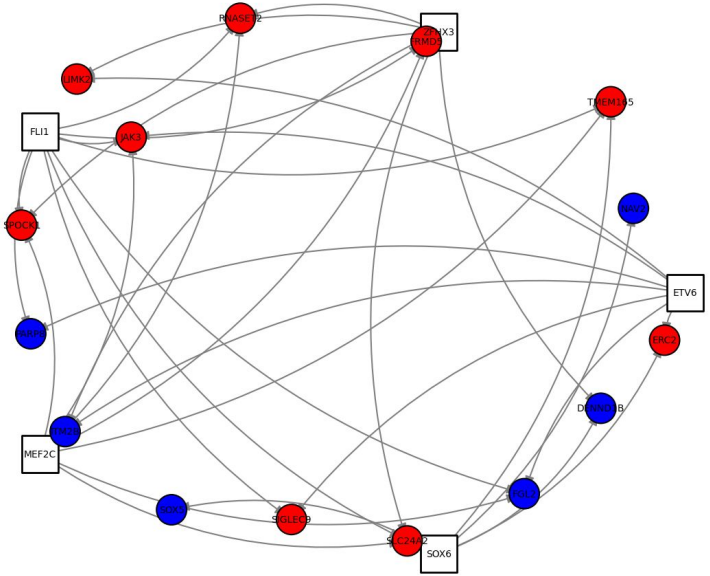
- Interferon Regulatory Factors (IRFs) are differentially regulated in HIV+ microglia
- Suggests interferon-driven immune response



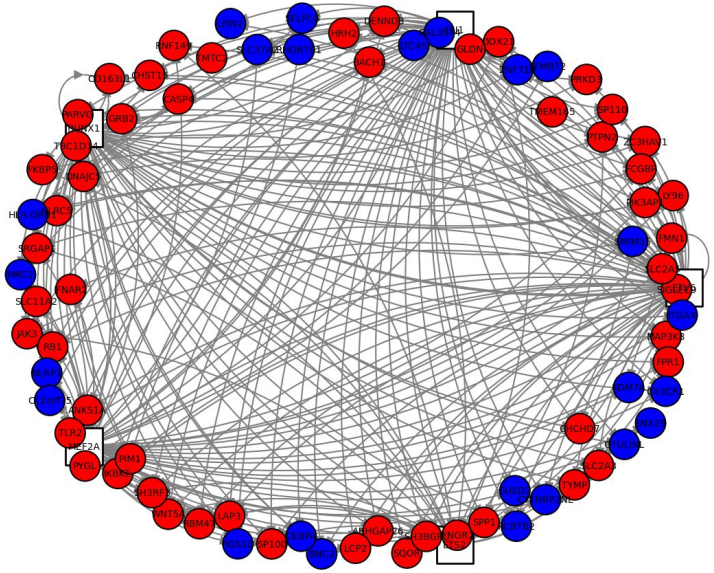
* Work in progress, not published

Co-regulated Target Network

Co-Regulated Microglia DEGs in PFC



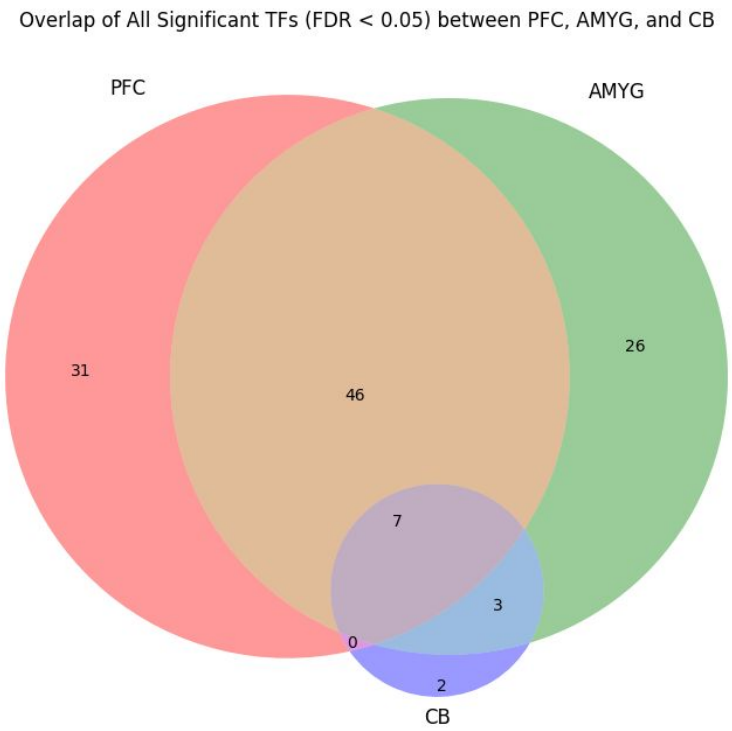
Co-Regulated Microglia DEGs in AMYG



* Work in progress, not published

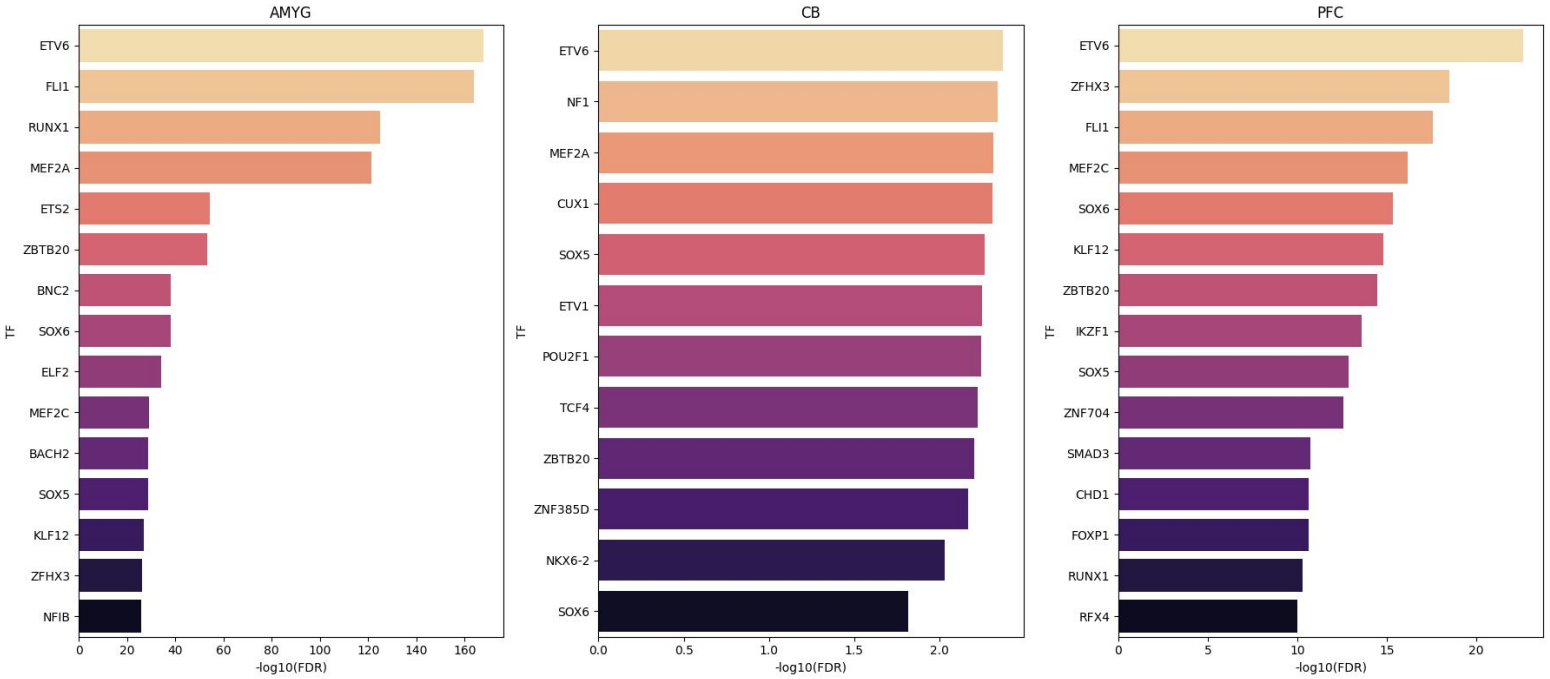
TF Enrichment Analysis

- Fisher's exact test identified 84 TF regulons significantly enriched in the HIV microglial DEG list (FDR < 0.05) in AMY, 82 in PFC, and 12 in CB.
- PFC and AMY show the most overlap
- Focus on mainly PFC and AMY



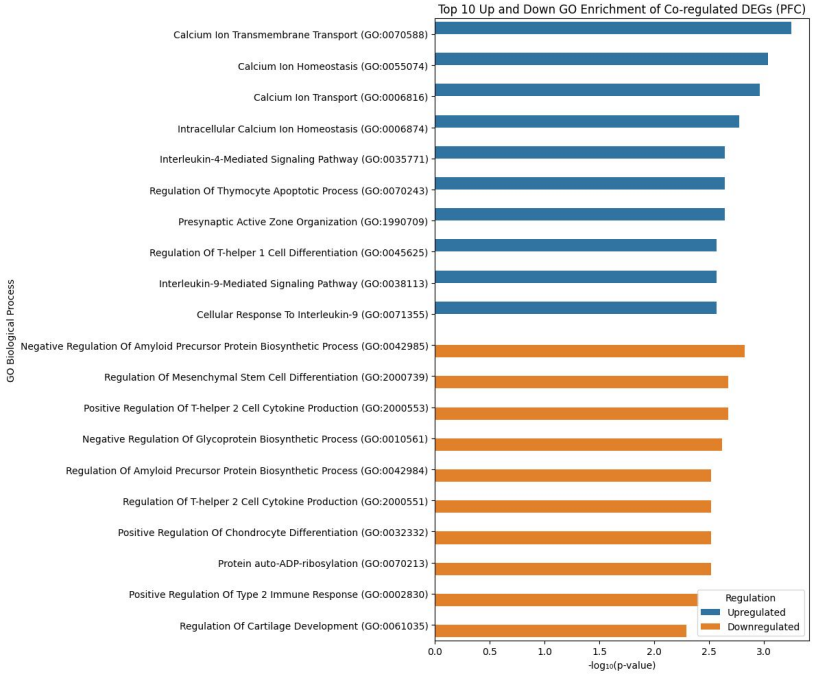
* Work in progress, not published

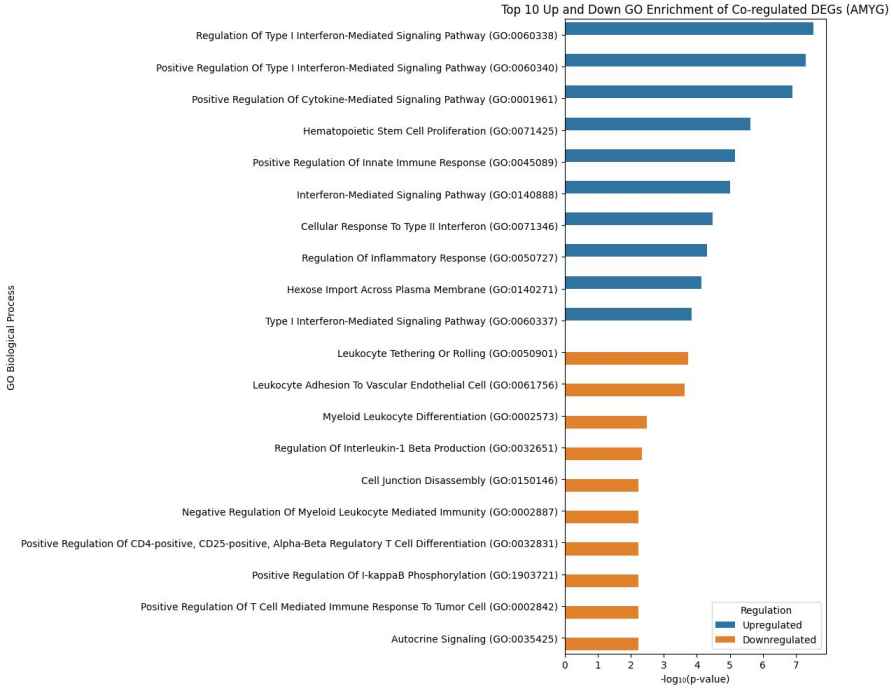
Top-15 Enriched TFs



* Work in progress, not published

Top 10 Up and Down GO Terms





* Work in progress, not published

Why Data Science Matters

- Biology now runs on big data
- Data science tools help us:
  - Clean and organize massive, noisy datasets
  - Build reproducible pipelines
  - Extract insights from multi-layered biology

Final Thoughts

Science is storytelling with data.
This is how we turn noisy cell measurements into real-world understanding of disease.

Want to learn more or help out? Let's talk!

Acknowledgements

Dr. Eran A. Mukamel

- Jason Chen
- Junhao Li, PhD
- Dominic Burrows
- Jo-Fan Chien, PhD
- Heng Xu
- Mike Cuoco
- Alon Gerber
- Rohini Gadde

Dr. Tariq Rana - Project Leader

- Shweta Jakhmola
- Shashi Tiwari
- Gulshanbir Baidwan
- William Bray

Dr. Kyle Gaulton

- Hannah Mummey
- Tanmayi Vashist
- Yuna Lee
- Parul Kudtarkar

Dr. Allen Wang, Center for Epigenomics

- Justin Buchanan
- Elizabeth Smoot
- Jacinta Lucero

Funding: *NIH Grant NIH/NIDA 1U01DA053630*

Section Materials

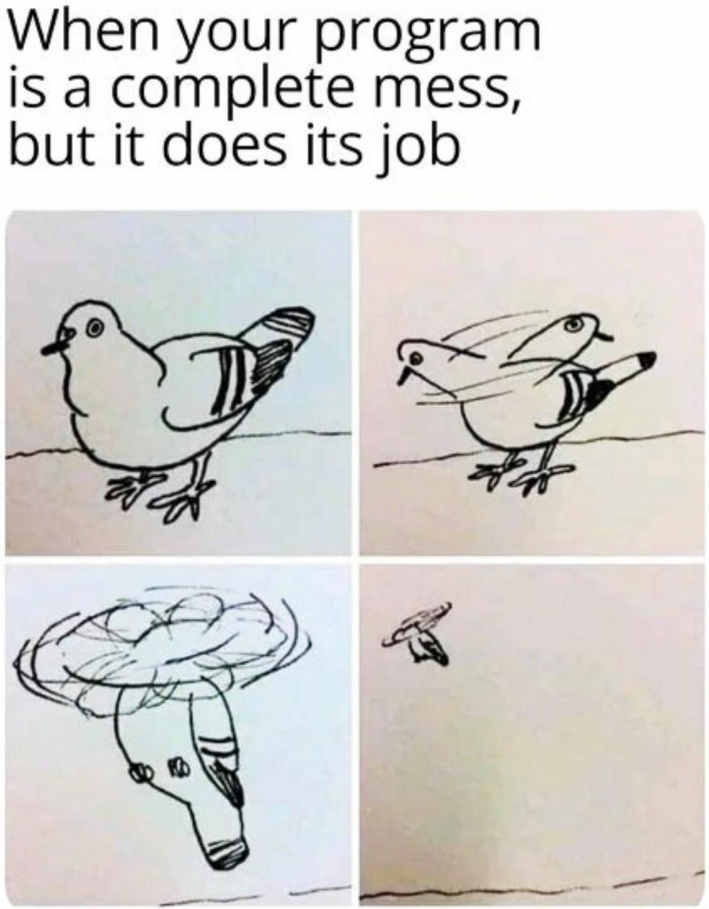
https://github.com/JasonC1217/COGS_108_B03-B04_Sp25/tree/master

or:

https://tinyurl.com/4d8wx3ne

SCAN ME

THANKS!

Questions on EdStem or office hours

Office hours: Thu, 3-5 PM