SINGLE CELL TRANSCRIPTOMICS IN SCHIZOPHRENIA POSTMORTEM BRAIN: MOVING BEYOND BULK LYSATE

Richard Crist
June 8th, 2020

Schizophrenia Transcriptomics

■ Microarray and RNA sequencing

■ Differentially expressed genes across many cortical and sub-cortical regions

- *Dorsolateral prefrontal cortex (dlPFC)* *(Fillman et al, 2013)*
- *Anterior Cingulate Cortex* *(Zhao et al, 2015; Hong et al, 2013)*
- *Superior temporal gyrus* *(Wu et al, 2012)*
- *Hippocampus* *(Hwang et al, 2013; Kohen et al, 2014)*
- *Amygdala* *(Chang et al, 2017)*

■ Enrichment of pathways and gene networks

- *Neural development*
- *Axon guidance*
- *Inflammation and immune-related proteins*

CommonMind Consortium

- Largest transcriptomic analysis of schizophrenia
  - *258 cases/279 controls*
  - *RNAseq in dlPFC*

- 693 differentially expressed genes

Fromer et al, 2016
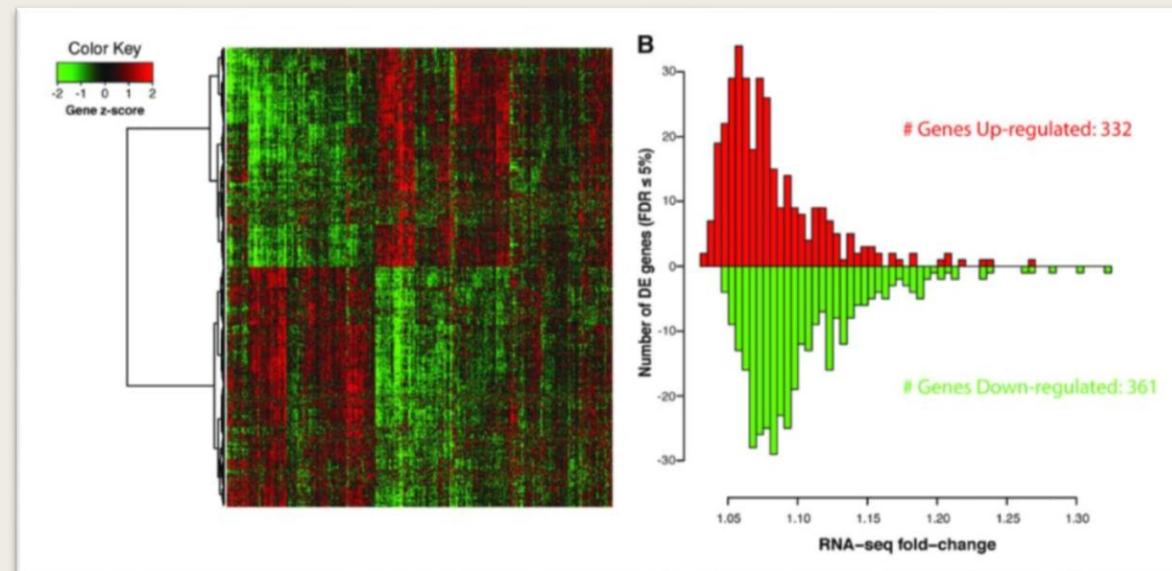
Cell Diversity in Postmortem Brain

Neurons
Astrocytes
Microglia
Oligodendrocytes
Pericytes
Vascular endothelial cells

- Brain, like all tissues, consists of many cell types
  - *Major cell populations (e.g. astrocytes)*
  - *Distinct sub-populations (e.g. PVALB+ interneurons)*

- Problems in assessing differential expression in bulk lysate
  - *Inability to identify which cells are affected*
  - *Missed expression changes in less common cell types*

Penney et al, 2019
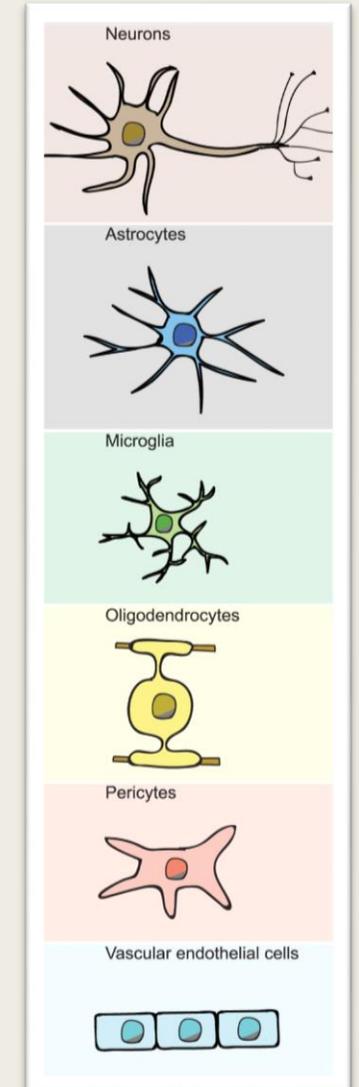
Schizophrenia Single Cell Transcriptomics

- Immunofluorescence and laser capture microdissection to collect individual populations of cells

- Layer III/V pyramidal neurons (Arion et al, 2017)
  - *72 PFC samples – 36 cases/36 controls*
  - *100 cells per layer for each sample*
  - *Expression assessed by microarray*
  - *1,783 differentially expressed probe sets corresponding to **1,420 genes***

- Parvalbumin positive (PVALB+) interneurons (Enwright et al, 2018)
  - *Same samples and methods*
  - *1,044 differentially-expressed probe sets corresponding to **872 genes***

Technical Hurdles with Single Cells

- Laser capture microdissection
  - *Low-throughput*
  - *Targeted*

- Pooling cells
  - *Lost information on variability between cells*
  - *Collapses sub-populations*

- Frozen human brain tissue
  - *Freeze/thaw ruptures cell membranes making single cell methods impossible*
  - *Single cells and nuclei have similar transcriptomes*
   - Same relative levels of **98% of transcripts** (Grindberg et al, 2013)

snRNAseq in Human Brain

REPORTS | GENETICS

Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain

Blue B. Lake[1,*], Rizi Ai[2,*], Gwendolyn E. Kaeser[3,4,*], Neeraj S. Salathia[5,*], Yun C. Yung[3], Rui Liu[1], Andre Wildberg[2], Derek Ga...

- Single nuclei isolated from frozen human brain
  - *FACS sort after staining for NeuN*
  - *One brain, multiple regions, multiple methods*

- Fluidigm C1 platform for capture and library preparation
  - *Microfluidics chip designed for single cell*
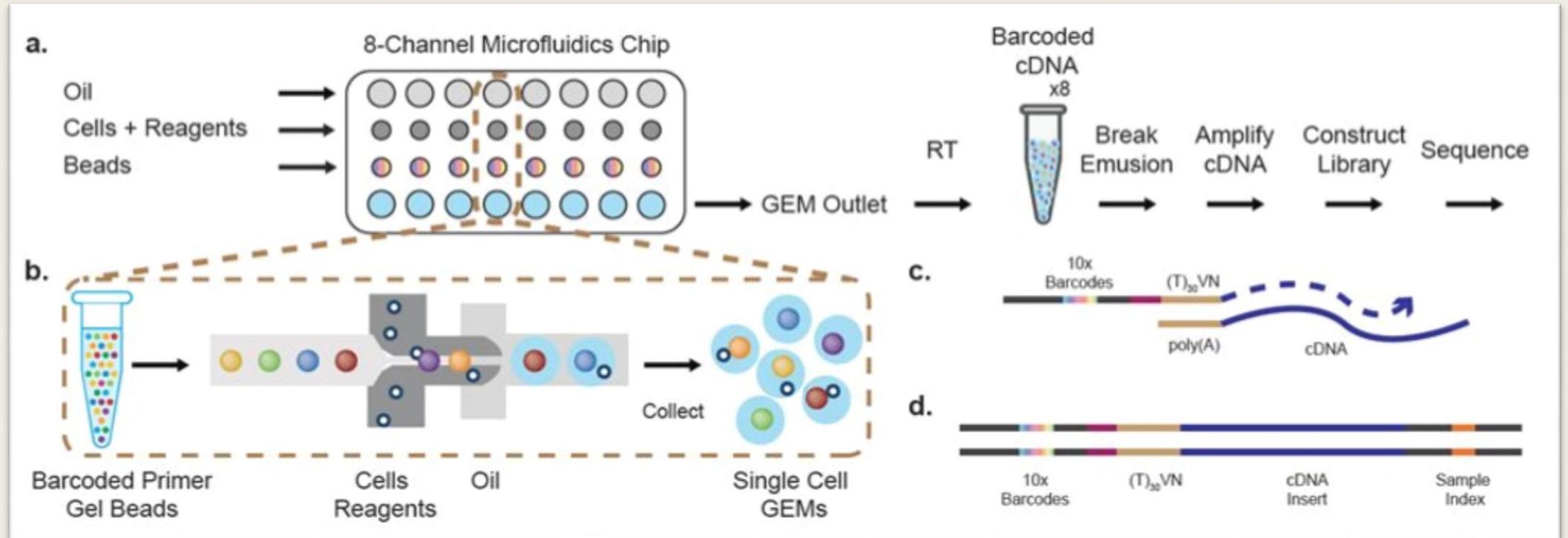 - *96 nuclei per chip*
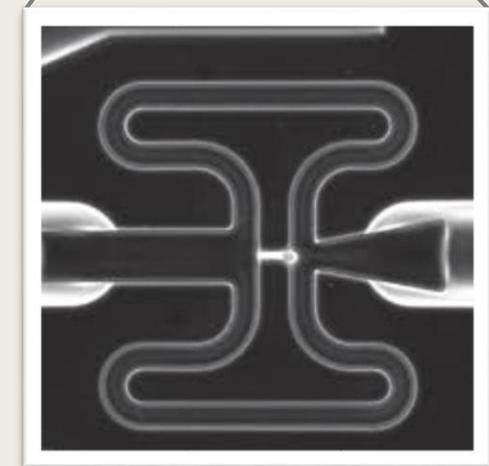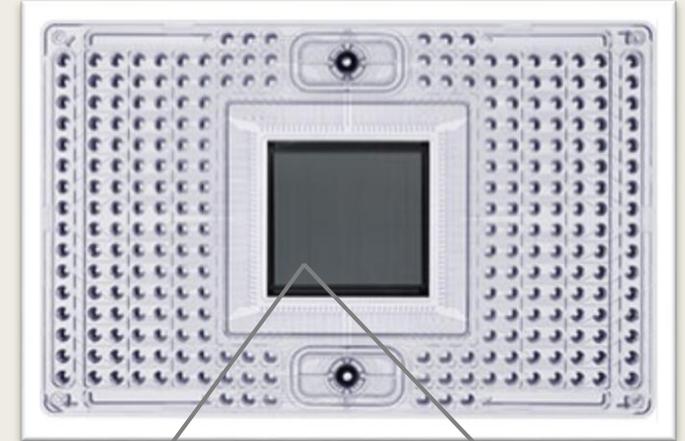
snRNAseq in Human Brain

■ Different layers of excitatory neurons

■ Distinct interneuron populations
  – *PVALB+, SST+, etc*

■ Regional differences
  – *e.g. Ex2 and Ex3 are layer 4 neurons from rostral and caudal areas, respectively*

Lake et al, 2015
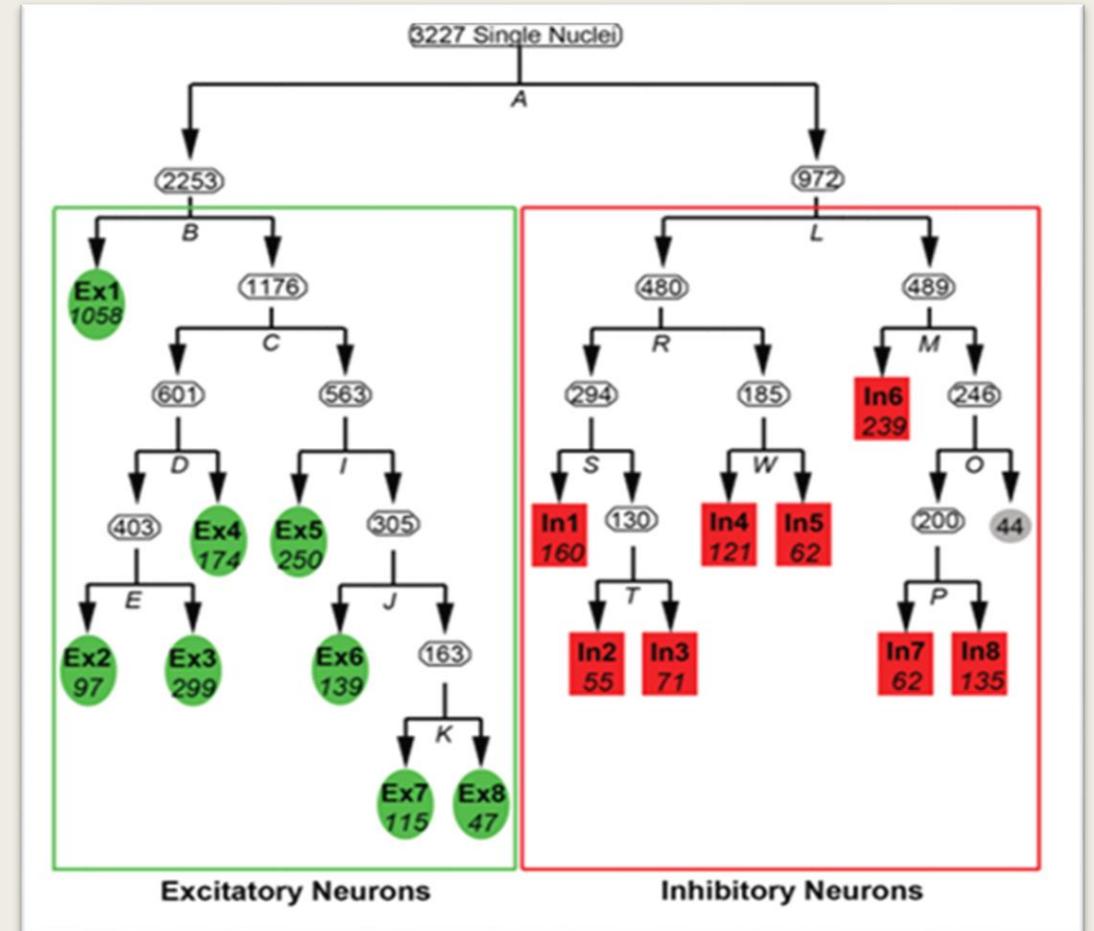
10x Genomics Chromium Controller

10x Genomics

- Pros
  - *High-throughput*
  - *Indexing at three levels: sample, nucleus/cell, transcript*
  - *Performs equally well on cells and nuclei*
  - *Theoretically cell type agnostic*

- Cons
  - *Expensive/fixed costs*
  - *Splicing not addressed (mostly)*
  - *Possibly not cell type agnostic?*

10x Genomics in Human Brain

- Five studies using 10x in postmortem human brain samples

  - *Alzheimer's* (Mathys et al, 2019)

  - *Autism* (Velmeshev et al, 2019)

  - *MS* (Schirmer et al, 2019)

  - *MDD* (Nagy et al, 2020)

  - *Huntington's* (Al-Dalahmah et al, 2020)

- Successfully distinguished cell populations using snRNAseq data

Experimental Design

- 32 postmortem dlPFC (BA9)
  - *16 cases (14M/2F)*
  - *16 controls (14M/2F)*
  - *European ancestry*

- Nuclei isolated using a modified version of previous protocol (Nagy et al 2020)
  - *Homogenization, 2x wash and filter, resuspend ~1,000 nuclei/μl*

- 10x library production and sequencing
  - *CHOP Center for Applied Genomics*

- Sequencing on NovaSeq 6000


Modified from Nagy et al., 2020
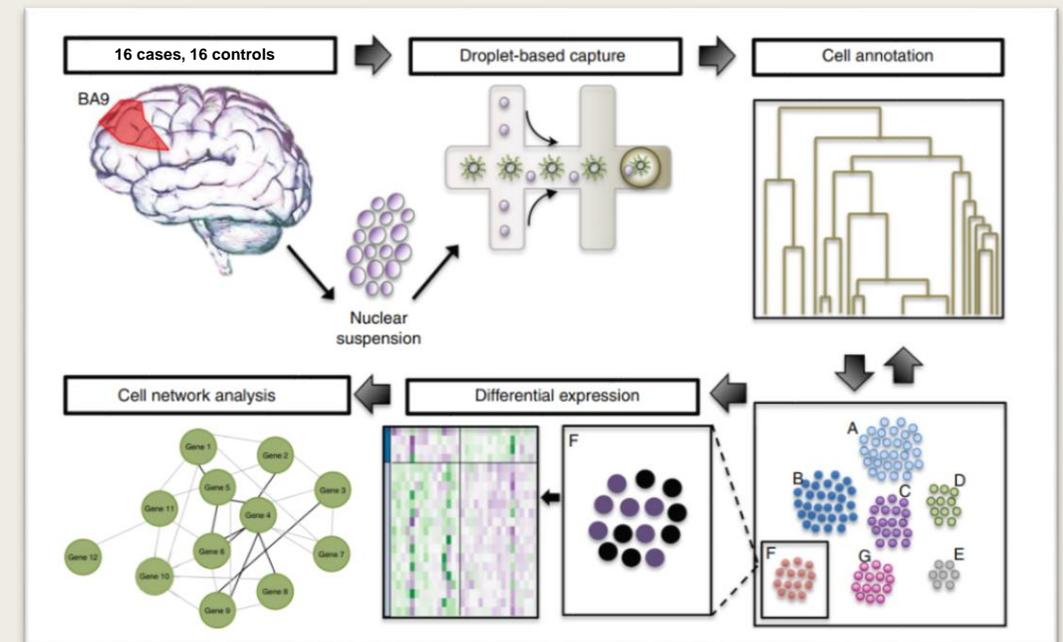
Quality Control and Clustering

- Deconvolution and alignment with 10x CellRanger

- QC using Seurat package
  - *Genes <3 nuclei, nuclei <200 genes, nuclei with high or low UMI*
  - *Normalize to 10,000 counts/nucleus*

- Clustering
  - *First 50 PCs calculated from 2215 "highly variable genes"*
  - *Low resolution first run, remove additional low UMI clusters and two SZ samples*
  - *High resolution second run, remove clusters specific to only a few individuals*

- Final results: ~323,821 nuclei in 27 clusters
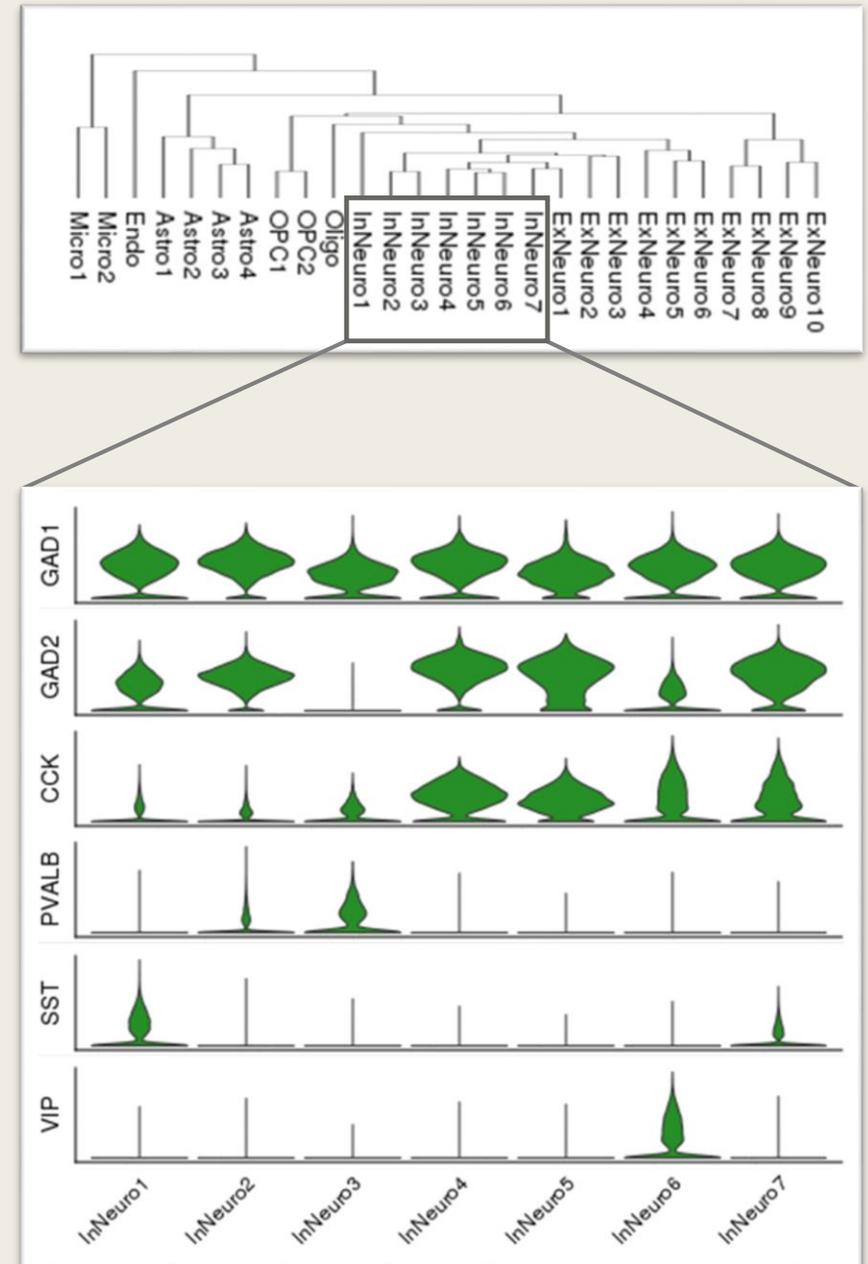 - *Previous 4 papers: ~313k total*
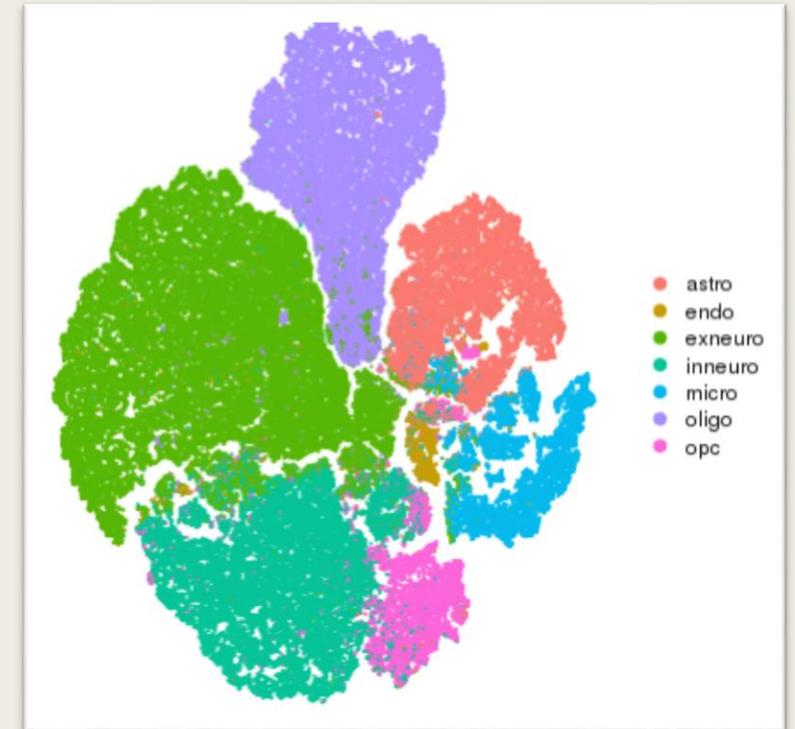
Cell Type Proportions
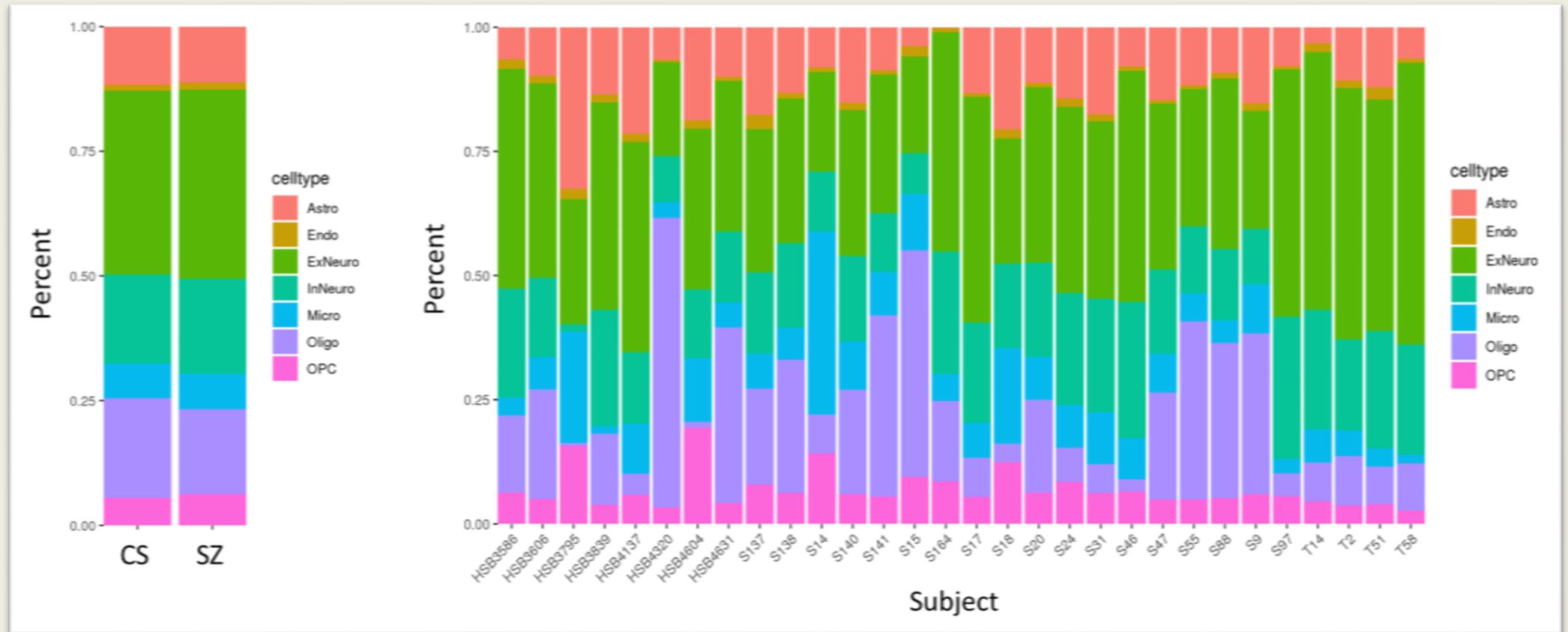
Sub-populations

Sub-populations

- Our single nuclei transcriptomic data differentiates known sub-populations
  - *Multiple interneuron types*
  - *Layer markers*

- How many clusters is the correct number?
  - *Sample size*
  - *Number of nuclei*
  - *Sequencing depth*

Expression Data in snRNAseq

- Single cell/nuclei RNAseq data forms a bimodal distribution for each gene
  - *Genes will not be detected in all cells or nuclei*
  - *Large number of zeroes*

- What is the best way to handle this in a statistical model?
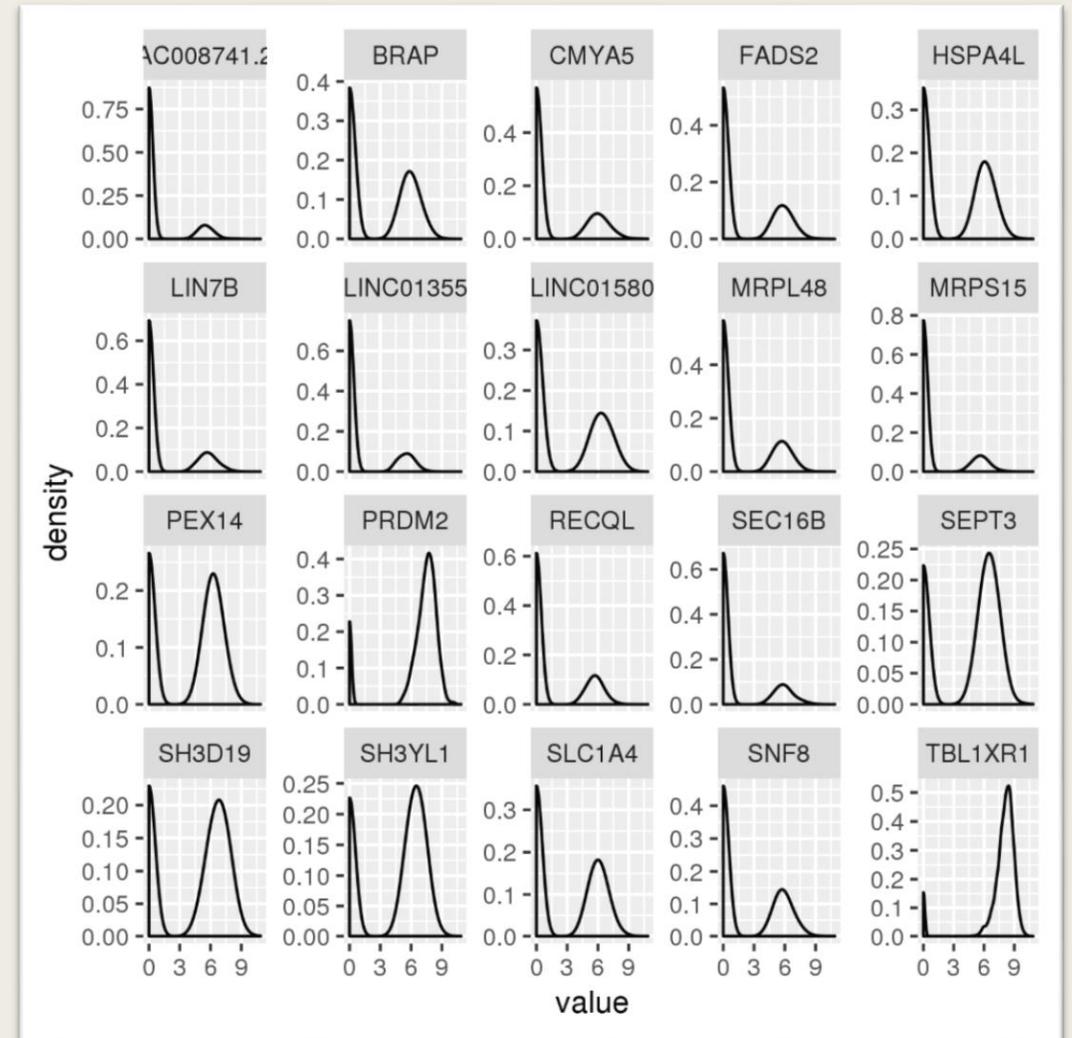
Analyzing snRNAseq: The Wild West

- There is no consensus yet on the best methods for analyzing snRNAseq data
  - *Mathys – LMM/Wilcoxon*
  - *Velmeshev/Schirmer – MAST*
  - *Nagy – LMM*

- The methods for those manuscripts in the field also include different:
  - *Covariates*
  - *FDR – 0.05 vs 0.1*
  - *logFC Cutoff – 0.25 vs 0.14 vs none*

Analyzing snRNAseq: The Wild West

- Systematic analysis of statistical methods
  - *Wide variation in efficacy*
  - *Single cell methods on average were not better than other methods*

- May be specific to the data set being analyzed or the questions being asked
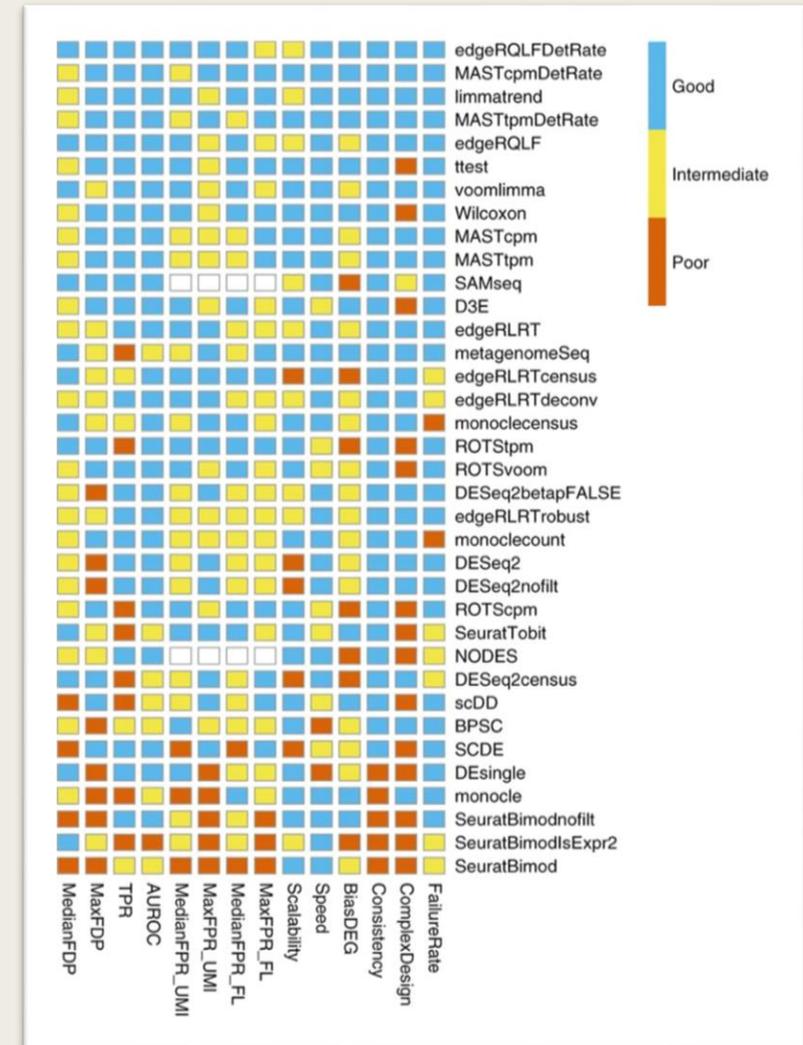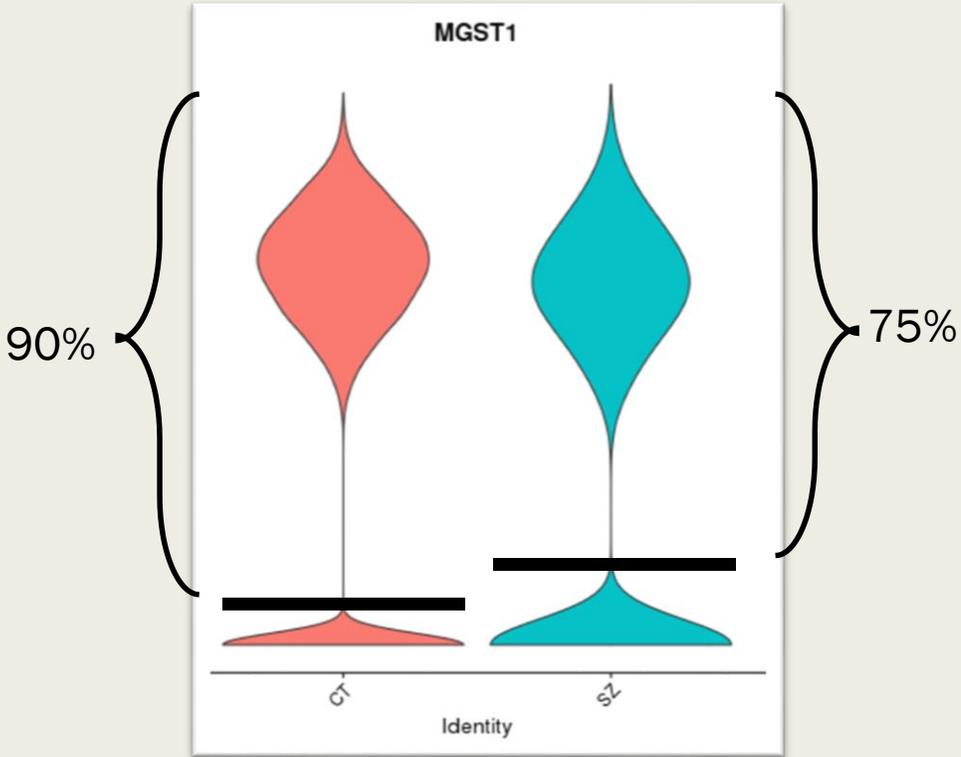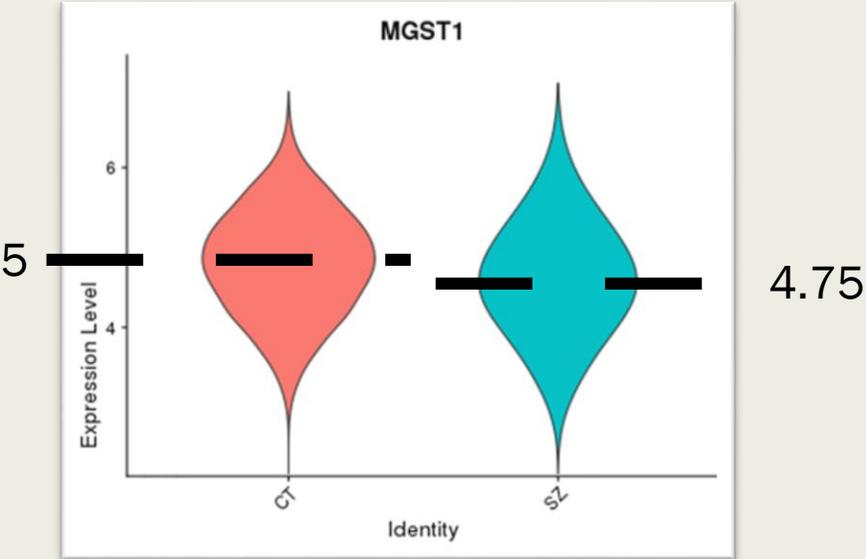
The *MAST* Hurdle Model
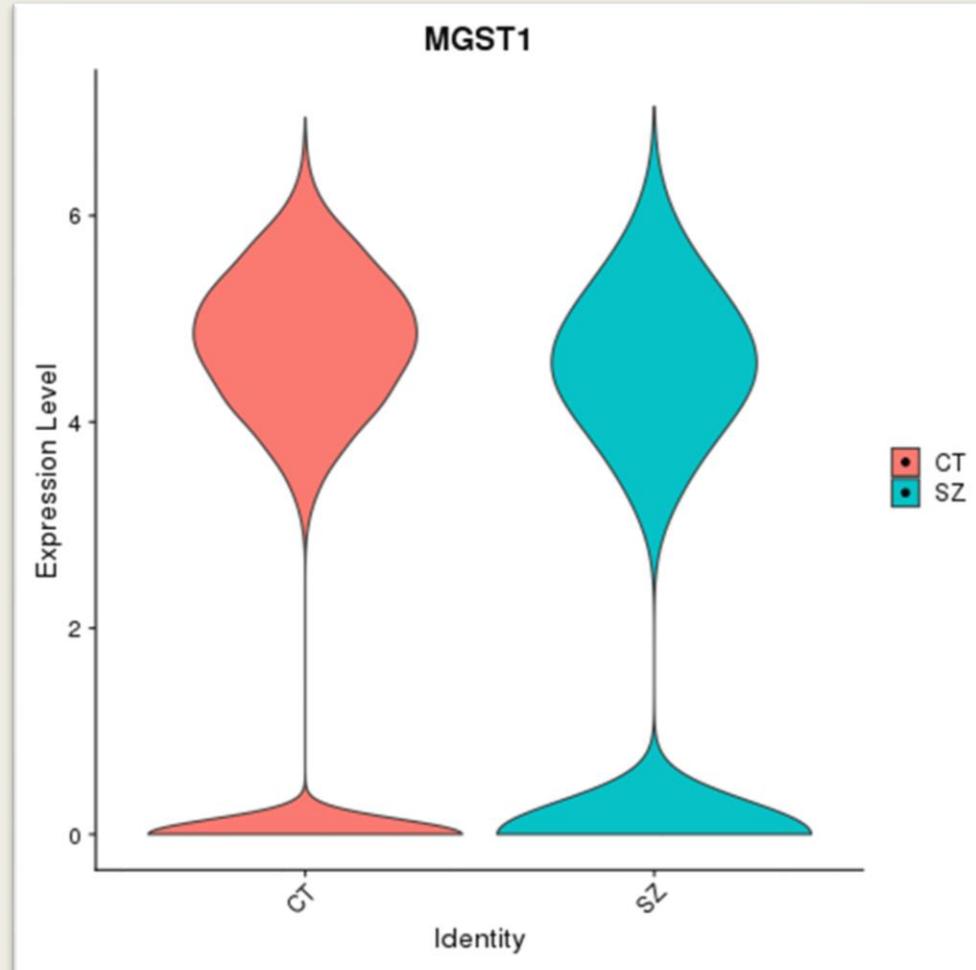
- Hurdle model is a combination of two different models to address the bimodal distribution of snRNAseq data

- Discrete model
  - *Is the gene detected in a larger percentage of nuclei in cases compared to controls?*

- Continuous model
  - *In non-zero nuclei, is expression of the gene higher in cases compared to controls?*
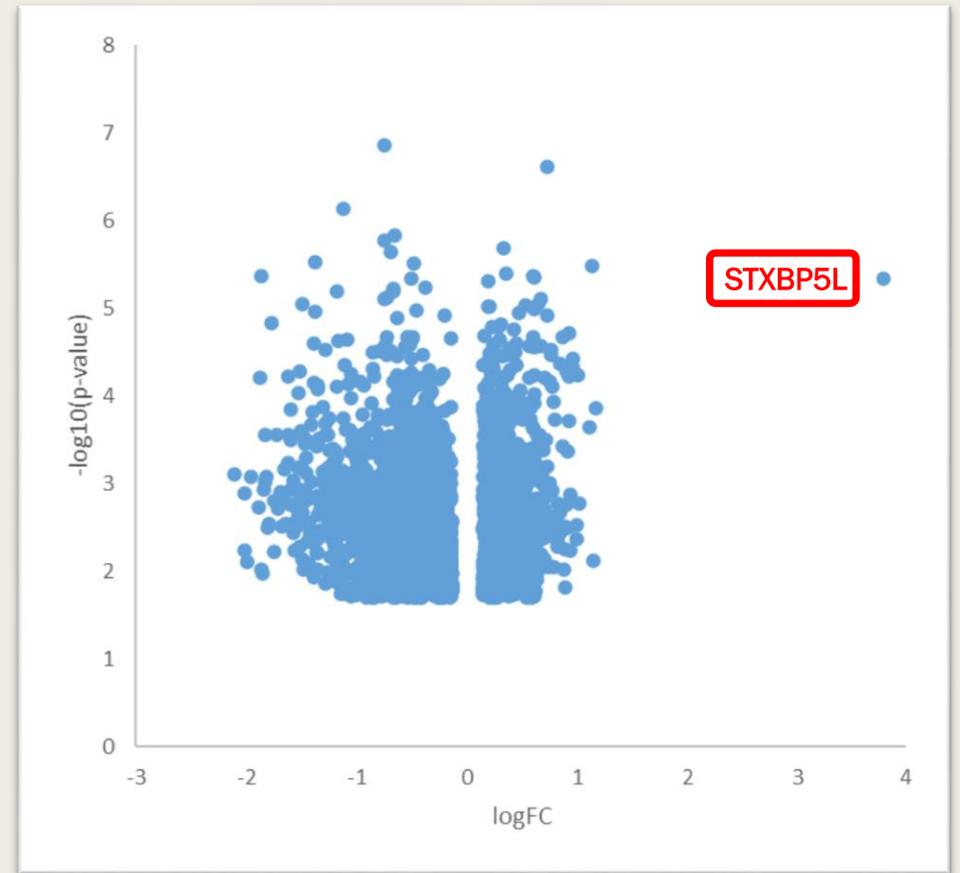
The *MAST* Hurdle Model

Discrete Model

90%

75%

Continuous Model

5

4.75

Differential Expression Analysis

- MAST – Hurdle Model

- Fixed effects
  - *Case/control status, sex, age, batch*
  - *Gene detection rate*

- Random effect
  - *Subject*

- Significance
  - FDR = 0.1
 - $Log_2FC \geq 0.14$ (10% difference)

Differentially Expressed Genes

- Differential genes found in 21/27 clusters

- 2,853 differentially expressed genes
  - *957 upregulated, 1896 downregulated*

- 2,196 unique genes

- Top 5 clusters accounted for 95.9% of differentially expressed genes

| Cell Type	# DEGs
Inhibitory Neuron #2 – PVALB+	1092
Excitatory Neuron #1 - Layer V HTR2C+	814
Excitatory Neuron #4 – Layer II/III	340
Inhibitory Neuron #5	254
Excitatory Neuron #3 – Layer IV	237
Excitatory Neuron #6 – Layer VI	53
Oligodendrocyte	18
Excitatory Neuron #9 – Layer V	7
Microglia #2	6
Excitatory Neuron #5 – Layer II/III	6
Astrocyte #4	5
Inhibitory Neuron #6 – VIP+	4
Excitatory Neuron #2 – Layer IV/V	4
Excitatory Neuron #10 – Layer VI	3
Inhibitory Neuron #4	2
Inhibitory Neuron #7	2
Inhibitory Neuron #3 – PVALB+	2
Endothelial	1
Astrocyte #3	1
Excitatory Neuron #7	1
Inhibitory Neuron #1 – SST+	1

Neurons and Schizophrenia

■ Schizophrenia GWAS hits are enriched for genes expressed in mouse neurons (Skene et al, 2018)

  – *Medium spiny neurons*
  – *Pyramidal cells in hippocampal CA1*
  – *Pyramidal cells in somatosensory cortex*
  – *Cortical interneurons*

■ Similar results from limited human data (Skene et al, 2018)

■ PVALB+ interneurons have reduced density or altered gene expression in schizophrenia (Chung et al. 2016; Enwright et al. 2018; Fung et al. 2014; Hashimoto et al. 2008; Joshi et al. 2015; Volk et al. 2016)

Skene et al, 2018
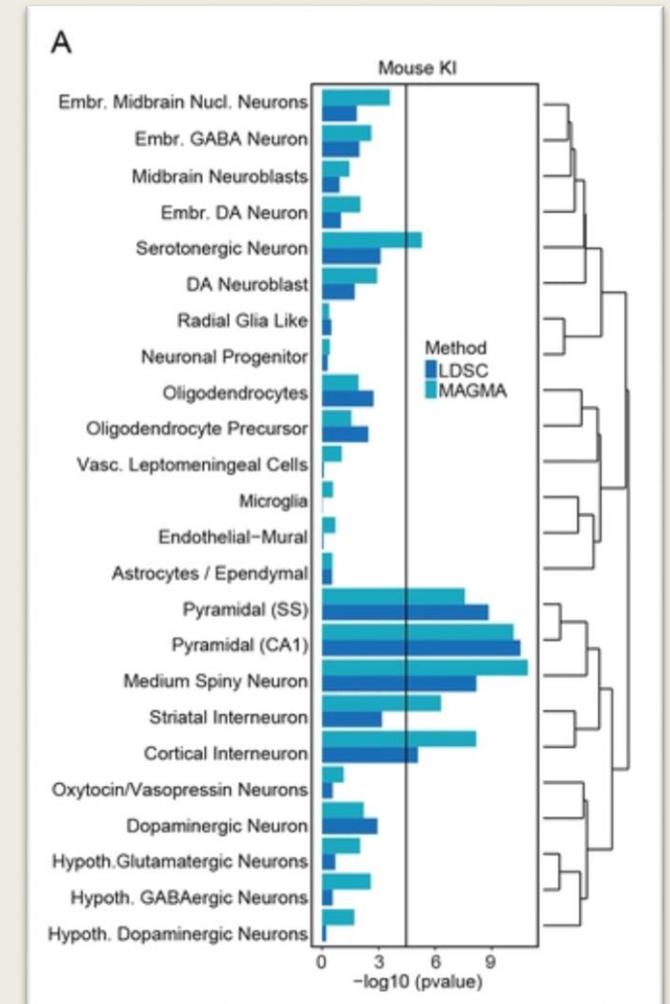
# Larger Clusters Do Not Have More Differentially Expressed Genes
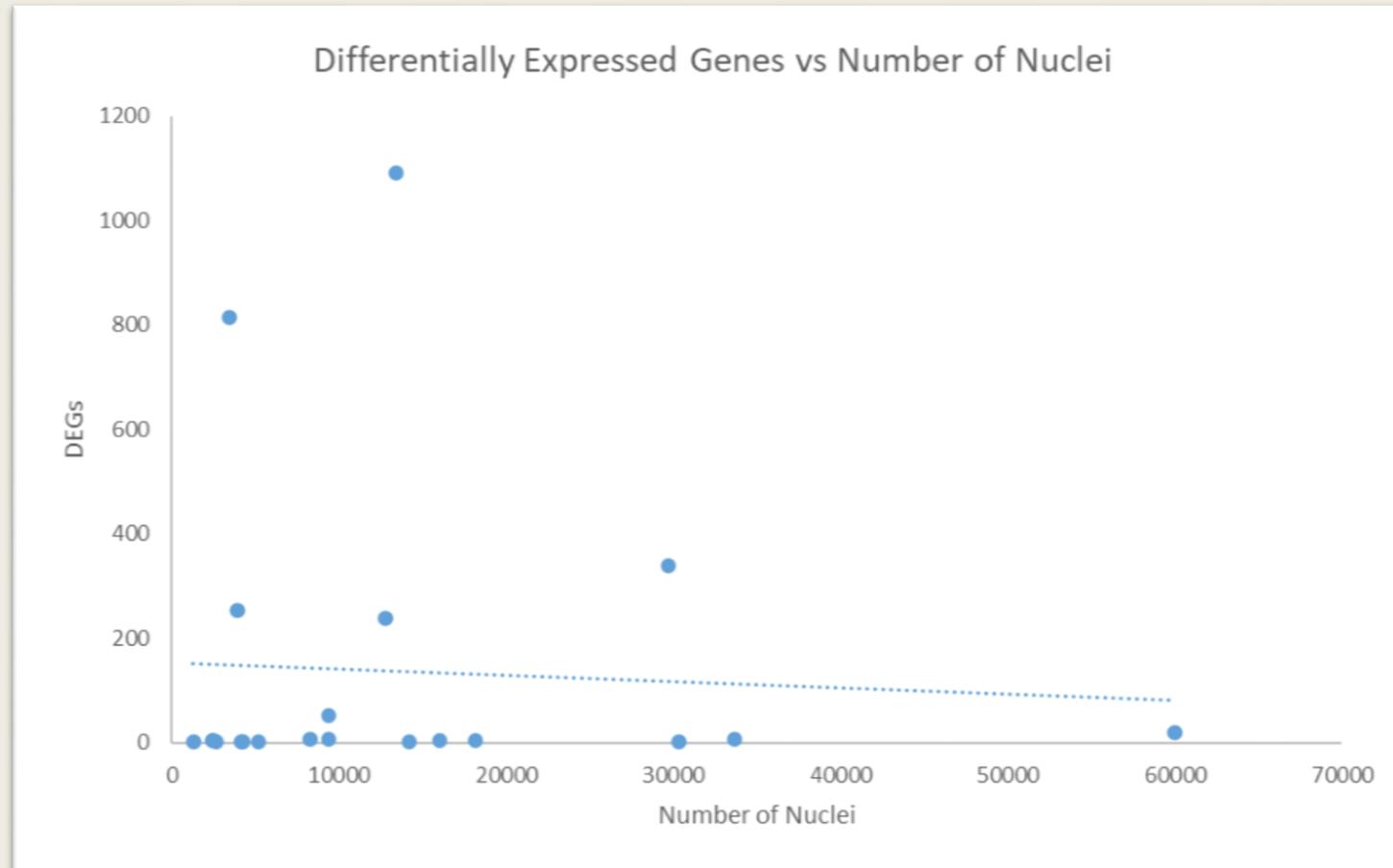
Cell Type Specificity
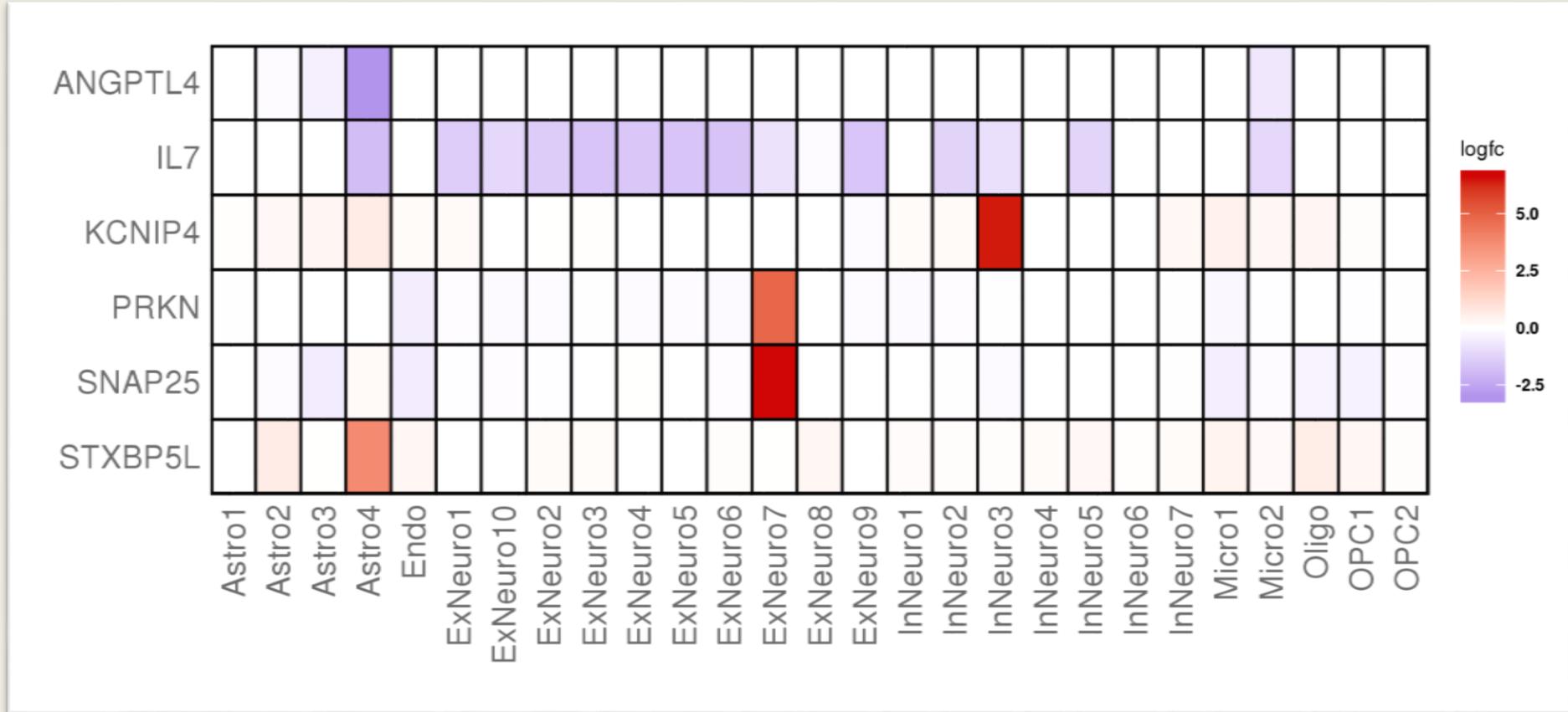
Molecular Clocks and Circadian Rhythm

- Schizophrenia patients frequently report sleep abnormalities (Kaskie et al, 2017)

- Evidence for altered rhythmic expression of clock genes in schizophrenia (Johansson et al, 2016)

- Inhibitory Neuron #2 – PVALB+
  - *Upregulated - CLOCK, CRY1, NPAS3*
  - *Downregulated - PER2, CSNK1D*

- Other Clusters
 - *CLOCK, ARNTL, BHLHE41*

Gooley et al, 2014
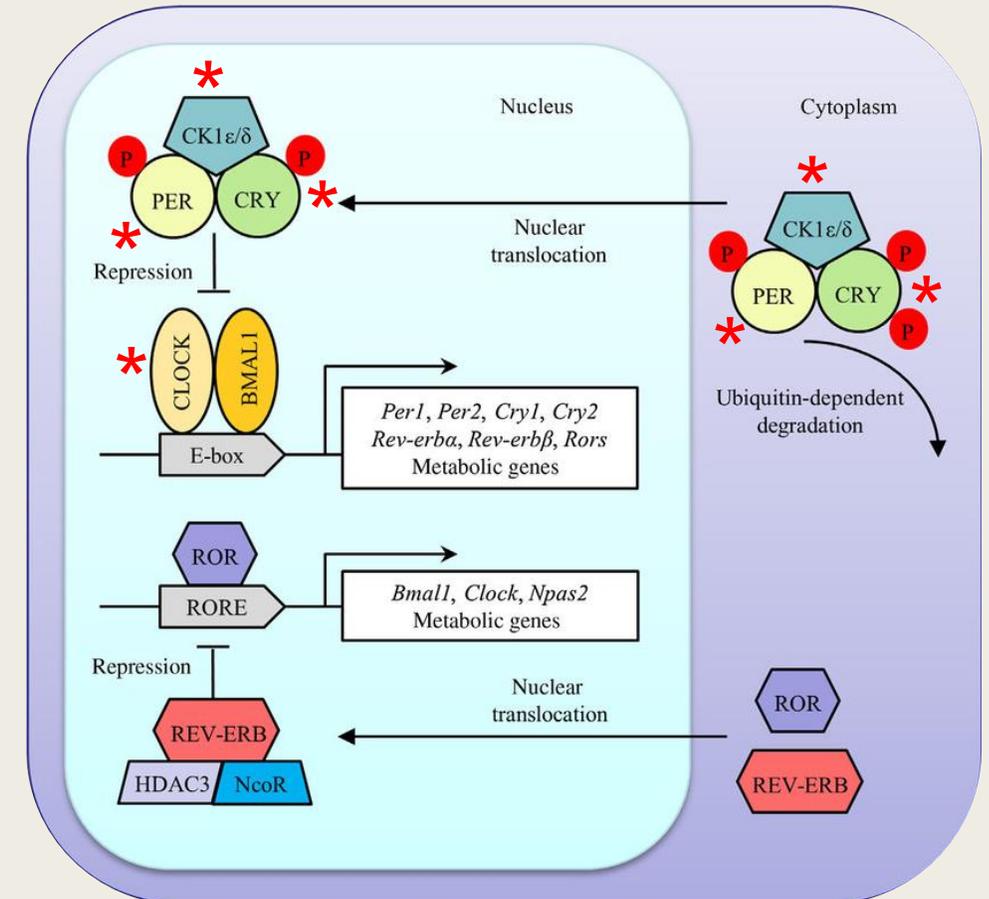
Schizophrenia as a "Channelopathy"

- Calcium channel genes
  - *CACNA1C*, *CACNA1L, CACNB1, CACNB3*

- Potassium channel genes
  - *KCNAB2, KCNC4, KCNJ3, KCNJ9, KCNK12, KCNK3, KCNQ5, KCNV1, KCTD2*

- Sodium channel genes
  - *SCN2B, SCN3B*

Ripke et al, 2014
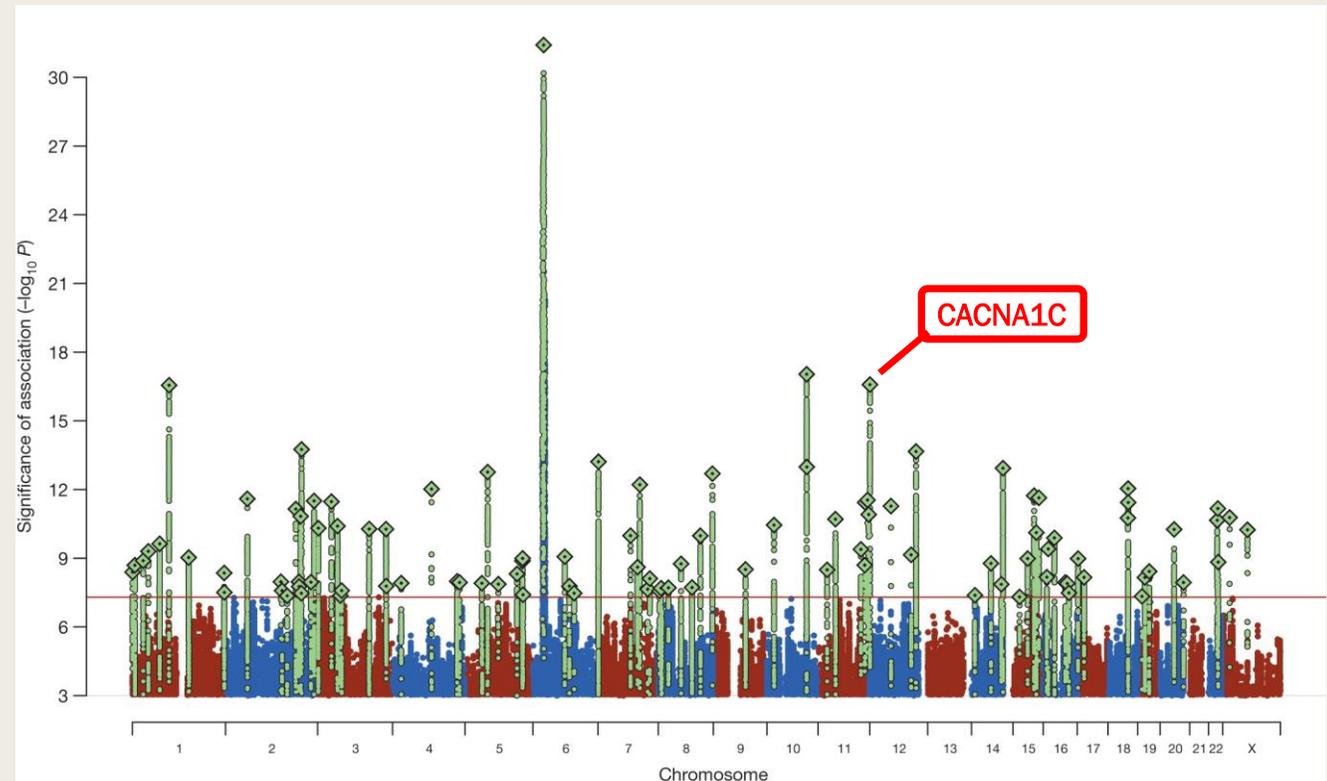
Gene Ontology Term Enrichment

■ Inhibitory Neuron #2 – PVALB+

– *Ubiquitin-Dependent Protein Catabolic Process*

– *Intracellular Transport*

– *Mitochondrion*

– *Spliceosomal snRNP Complex*

■ Excitatory Neuron #1 - Layer V HTR2C+

– *ATP Synthesis Coupled Electron/Protein Transport*

– *Ribonucleoprotein Complex Assembly*

– *Vesicle Fusion to Plasma Membrane*


Bousman et al, 2019
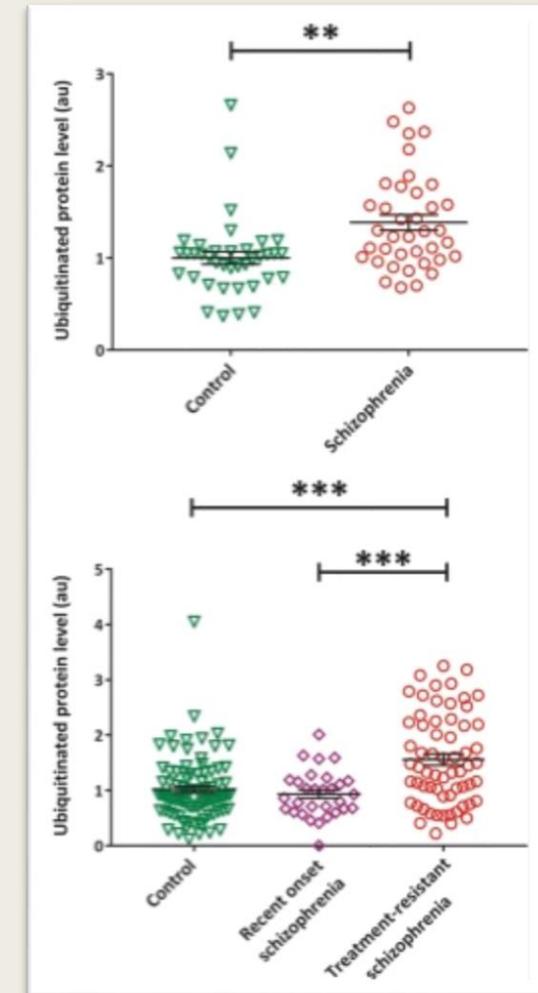
Mitochondria and Schizophrenia

- *NDUFV1 downregulated in five clusters*
  - *Previously linked to schizophrenia*

- Genes encoding mitochondrial proteins
  - *NDUFAB1, NDUFAF5, NDUFAF7, NDUFB2, NDUFB5, NDUFB7, NDUFB9, NDUFC1, NDUFS, NDUFS2, NDUFS5, NDUFS8*
  - *UQCC1, UQCC2, UQCR10, UQCRC1, UQCRC2, UQCRH, UQCRQ*
  - *COX18, COX20, COX4I1, COX5B, COX6B1, COX7B, COX7C*
 - *MTFR1*

Electron Transport Chain
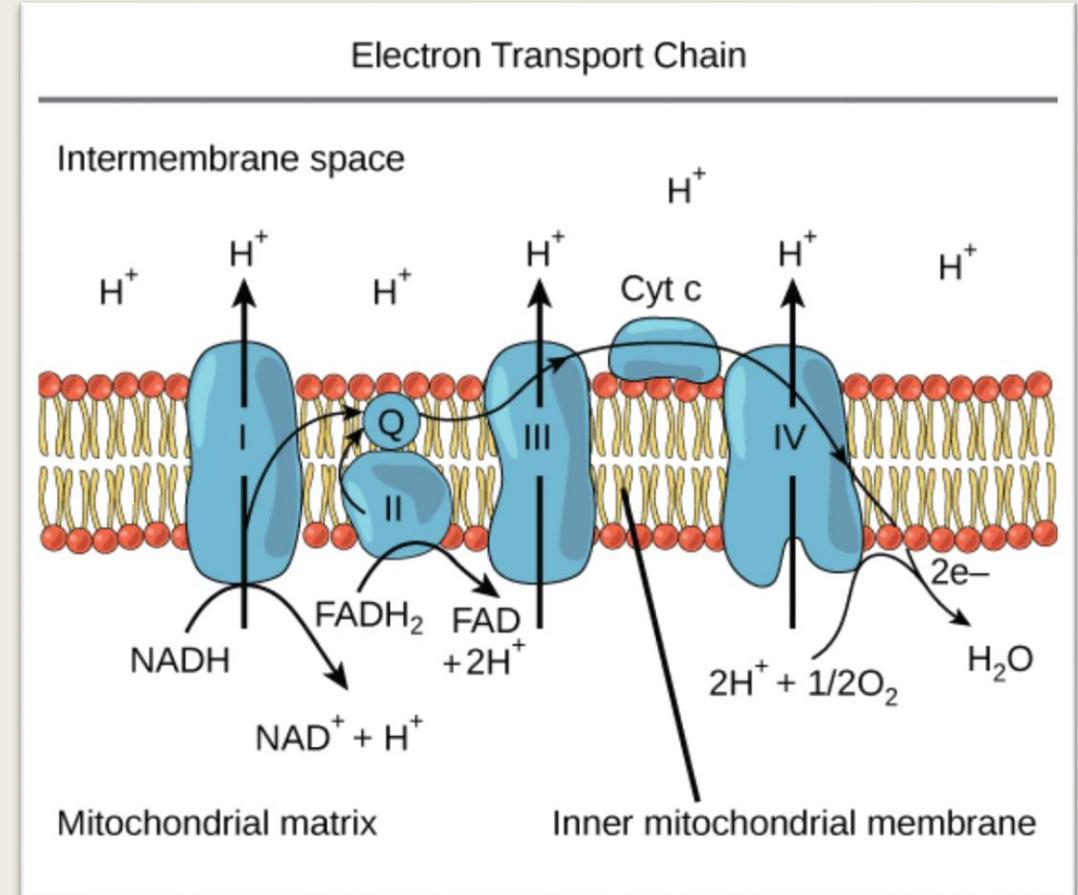
Ingenuity Pathway Analysis
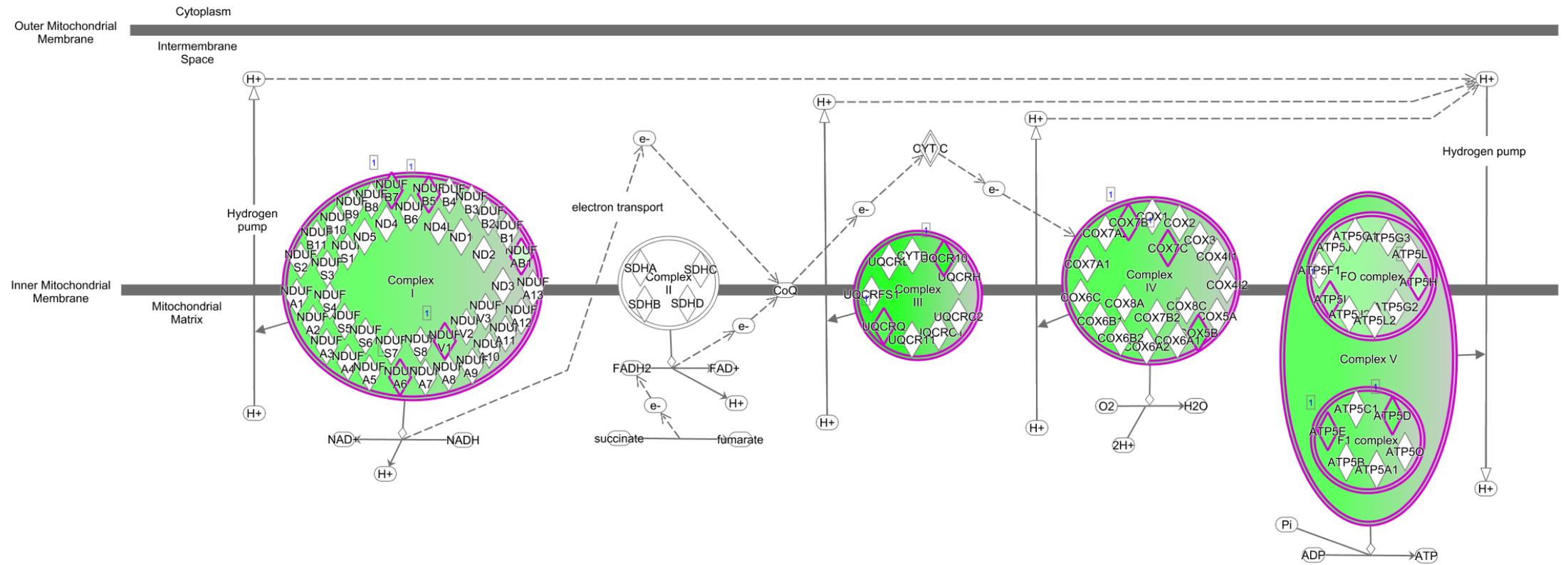
Transcription Factor Analysis

- A number of transcription factors appear in our list of differentially expressed genes
  - *Do the genes they regulate also appear?*
  - *Are those target genes enriched for particularly GO terms or pathways?*

- Target genes identified using TF2DNA

- Enrichment analyzed using DAVID
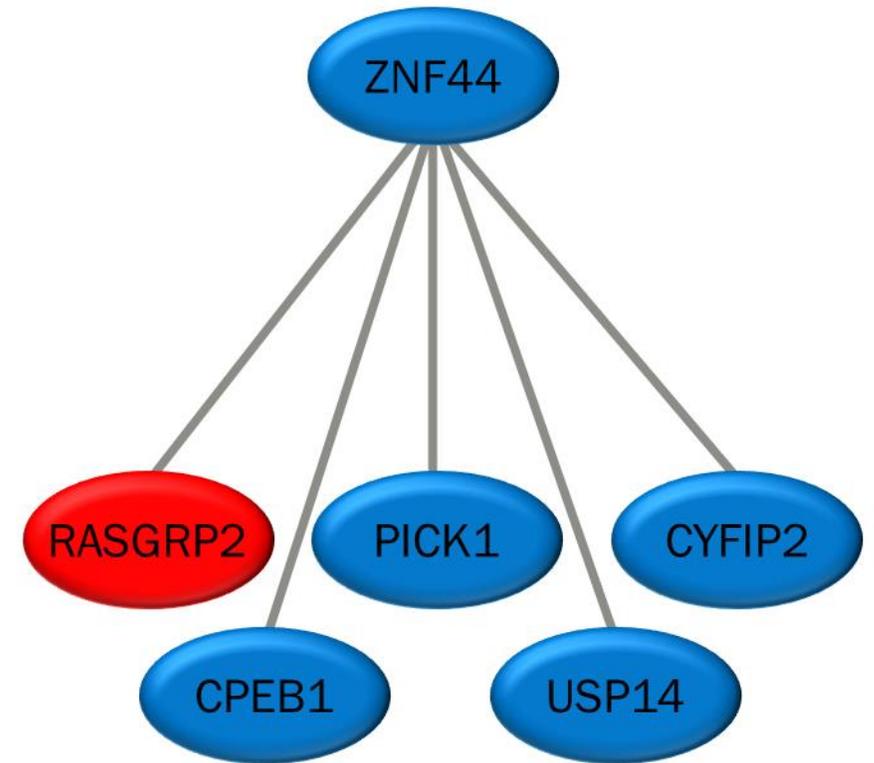
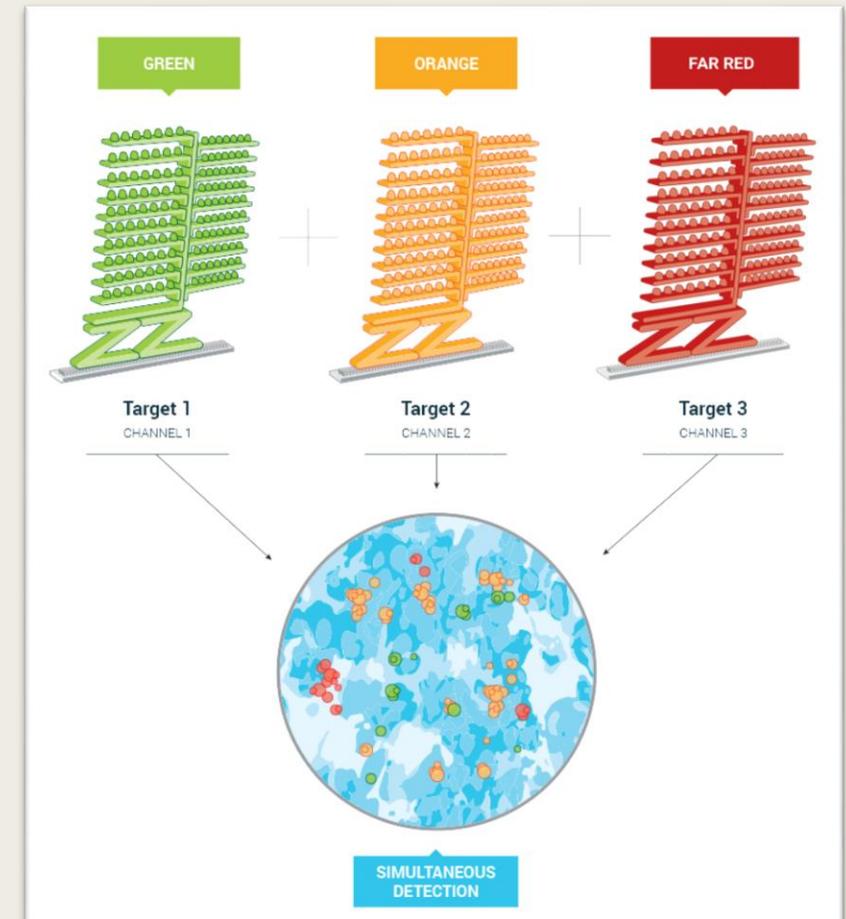

GO Cellular Compartment: Synapse

The Missing Glia

- Gray matter from the cortex contains more glia than neurons
  - *1.48 ratio* (Azevedo et al, 2009)

- snRNAseq from postmortem human brain has found the reverse

- Technical issues
  - *Nuclear isolation*
  - *Glia have fewer UMIs than neurons*

Mathys et al, 2019
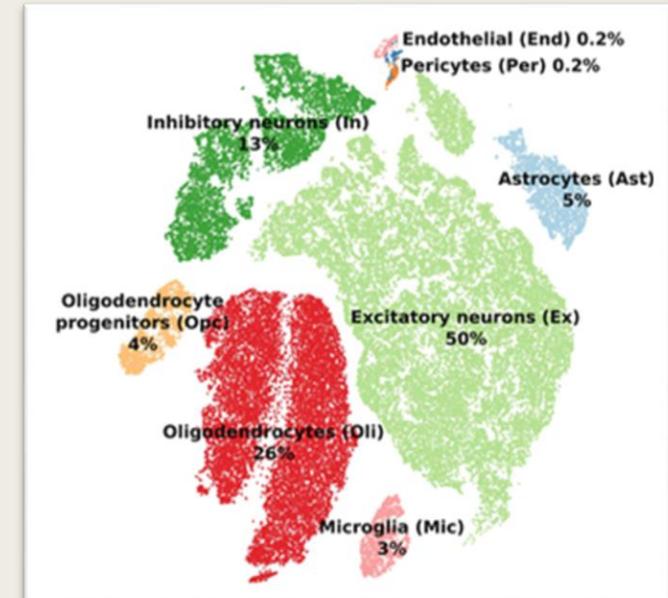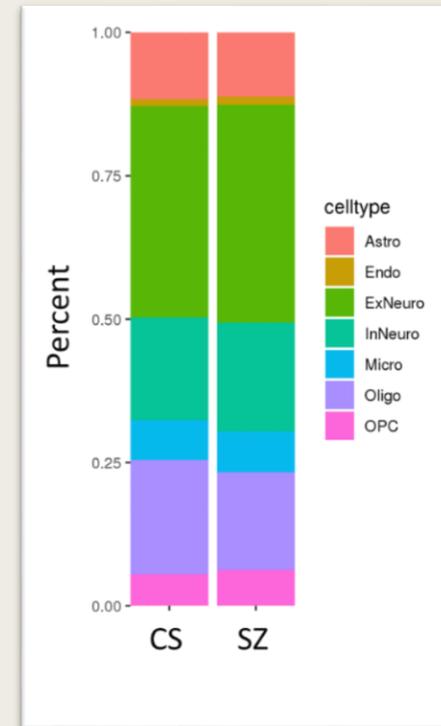
Future Directions

■ Non-sequencing Validation

    – *RNAscope with Dr. Lauren Stein*

■ Replication

    – *Sample size and individual variation*

■ Expand Into Substance Use Disorders

    – *Opioid use disorder*

    – *Animal models*

ACD Bio

Acknowledgements