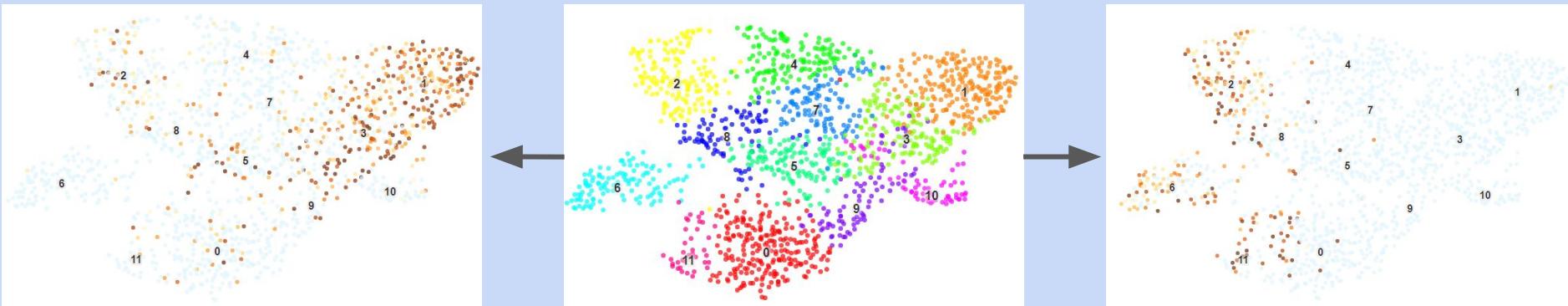


Understanding Developing Neuronal Subtypes within the Superior Colliculus

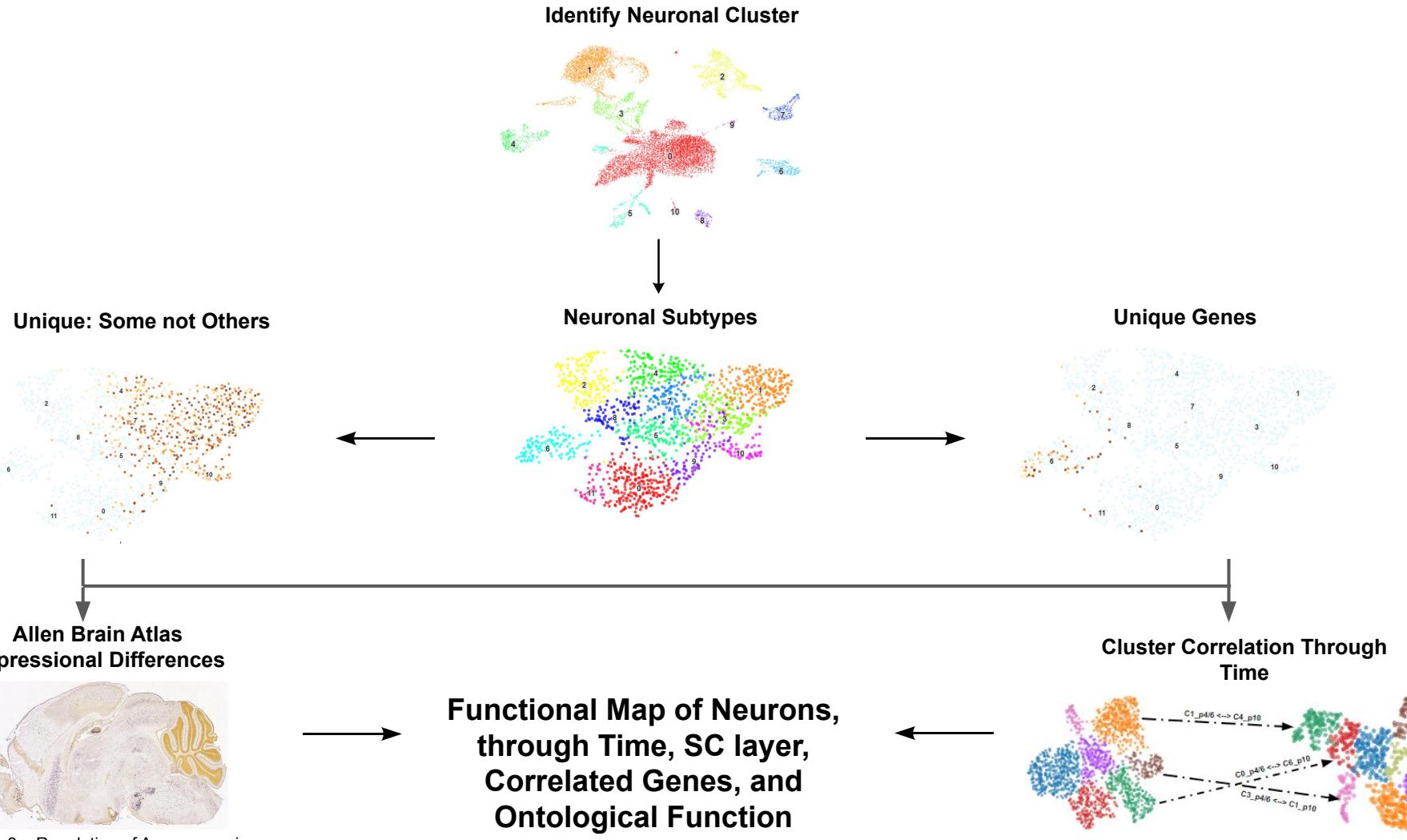


Background

- The neuronal subtypes within the SC remain relatively unknown
- As such, we want to categorize the various types of neuronal cells and see how they differentiate through time
- Lastly, we want to see how these neuronal subtypes are being expressed within different layers within the SC and how those change through time
- By categorizing the subtypes of neurons, seeing how they differentiate through time, and finding the concentration of subtypes, we can better understand how these neurons play a larger functional role in the SC

Project Overview: Goals and Objective

1. Identify the Neuronal Clusters
2. Identify Unique Genes for Individual and Collections of Neuronal Clusters
3. Identify all Possible Neuronal Subtypes
4. Find Functional Names for Genes using Ontologies
5. Find Expressional Differences between Genes using Brain Sectioned Images
6. Integrate Timepoints and Find Clustural Similarities



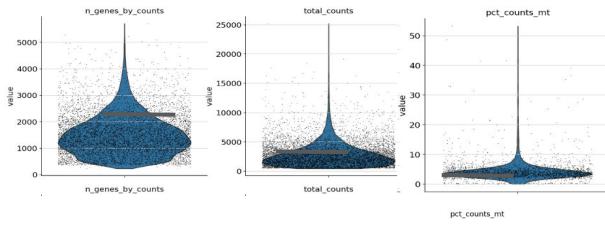
Project Overview: Methods

Scripts

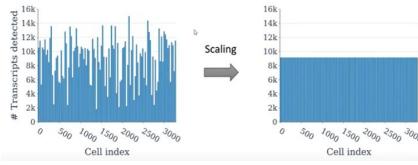
- Unique Genes in Clusters
- Unique Genes in Some but not All Clusters
- Allen/Panther/Cell Browser API
- Timepoint Integration (Harmony)
- Correlating Unique Genes
- Sankey Plots

scRNA-Seq: Methods Preprocessing/Filtering

Filtering Cells: Removing Outliers



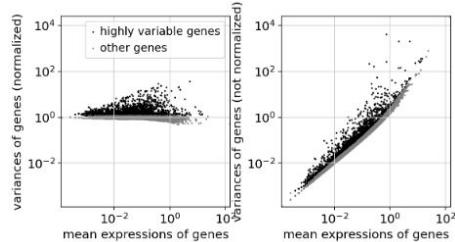
Scaling Data



Log Transform Data

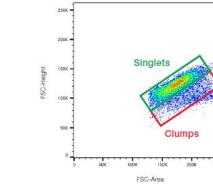
	Raw data			Log ₂ transform		
	Cell Type A	Cell Type B	Δ	Cell Type A	Cell Type B	Δ
Gene 1	1	2	1	0	1	1
Gene 2	100	200	100	6.64	7.64	1

Select Highly Variable Genes

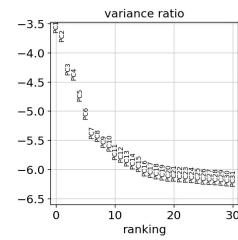
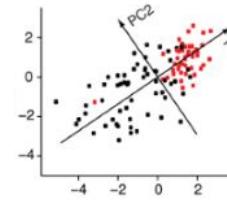


Pre-Processing/ Filtering Overarching Overview

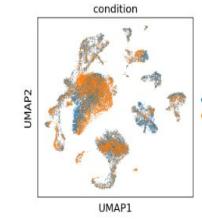
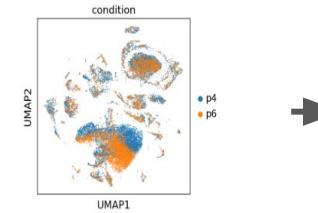
Doublet Filtering



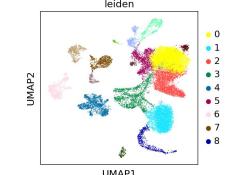
PCA



Batch Effect Correction

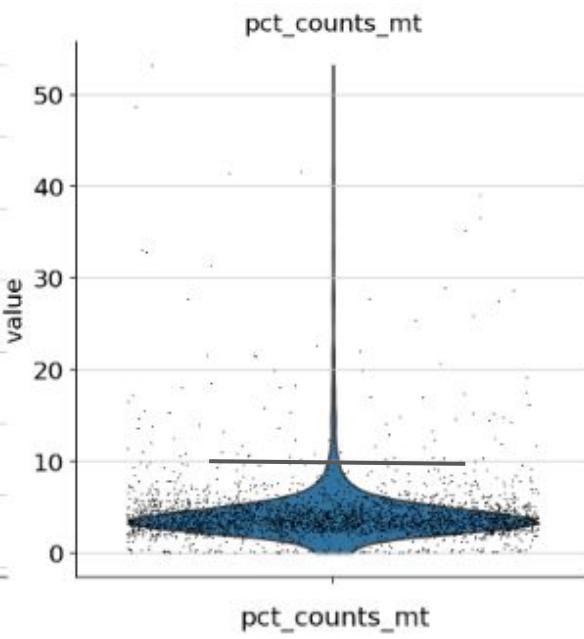
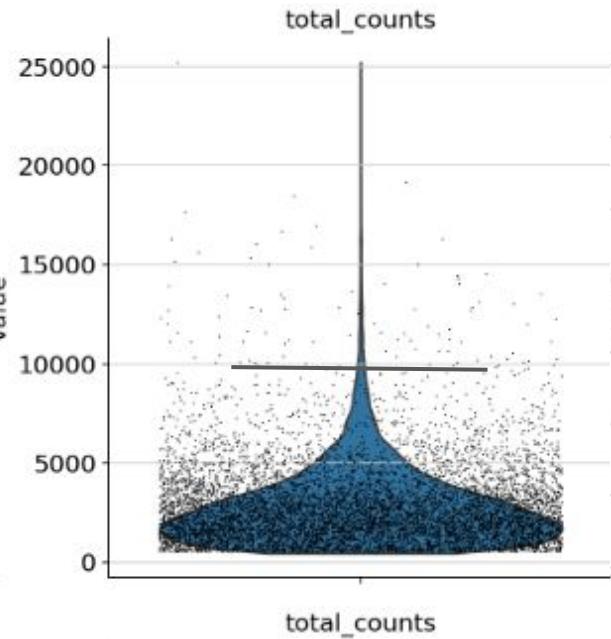
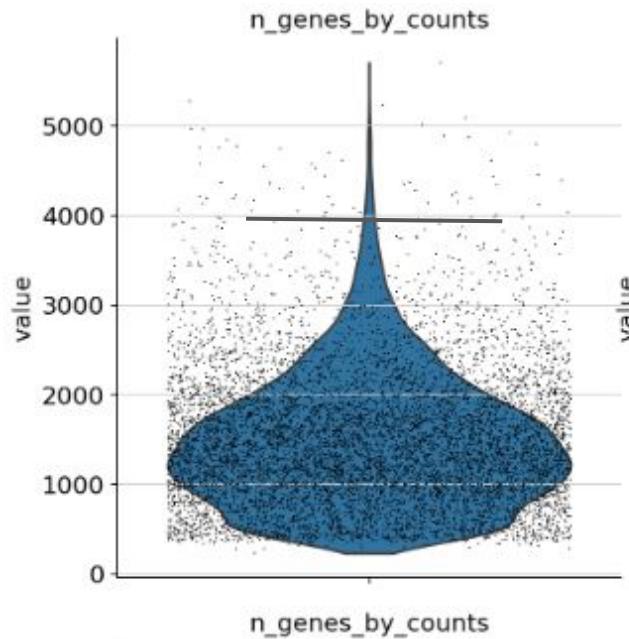


Clustering

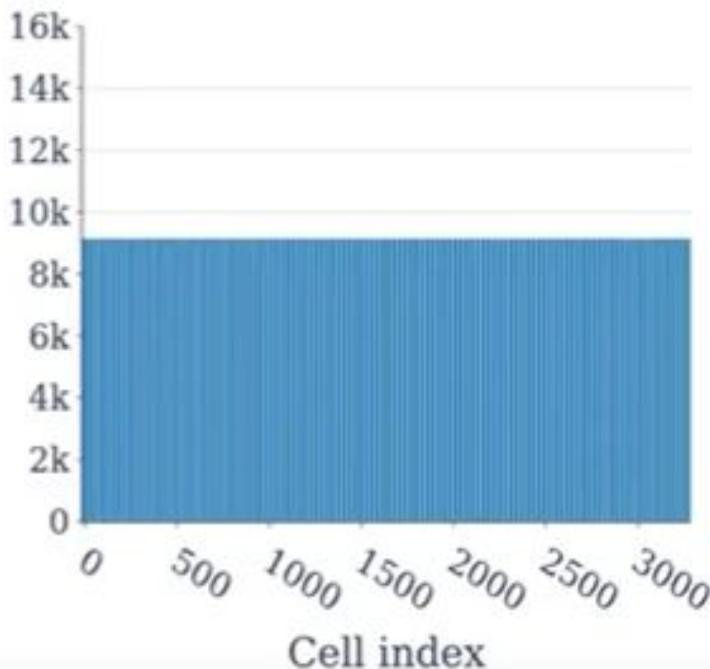
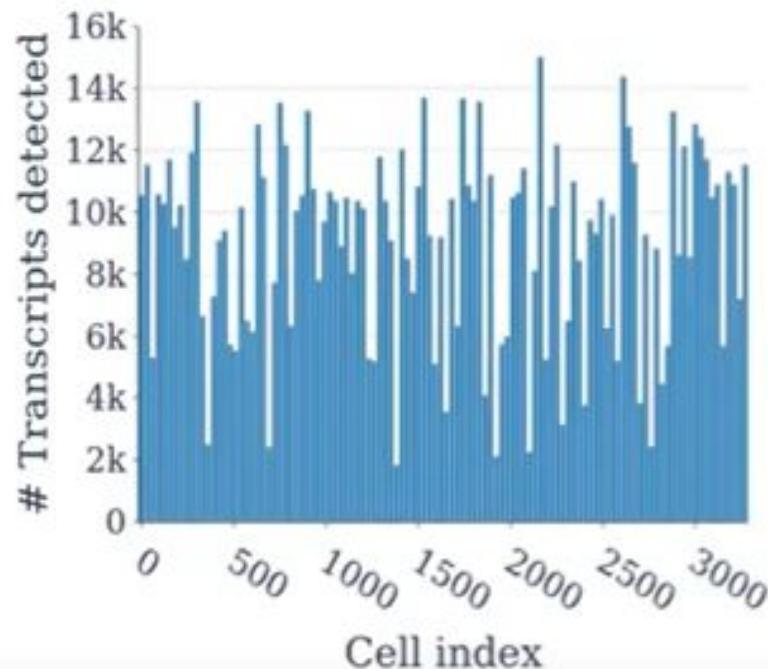


Filtering Cells: Removing Outliers

- Cutoff at 4000, 10000, and 10 counts respectively



Normalization: Scaling Transcript Counts



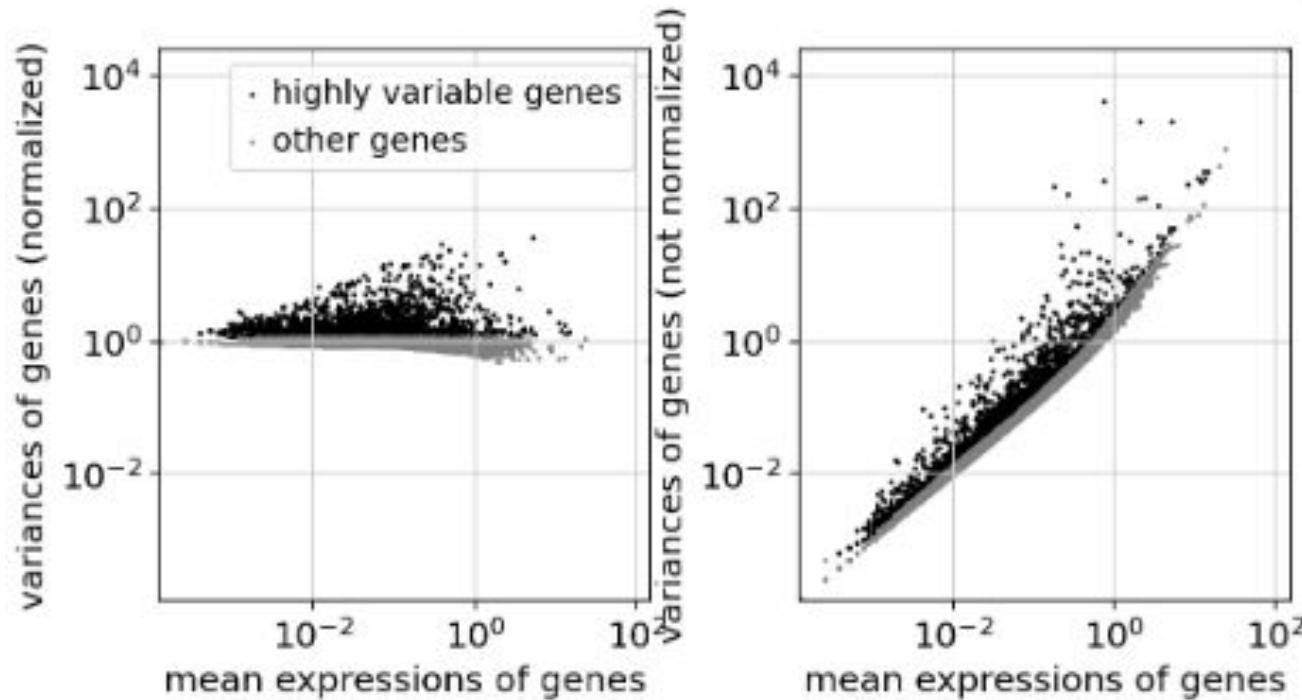
Normalization: Log Transform

	Raw data			Log ₂ transform		
	Cell Type A	Cell Type B	Δ	Cell Type A	Cell Type B	Δ
Gene 1	1	2	1	0	1	1
Gene 2	100	200	100	6.64	7.64	1

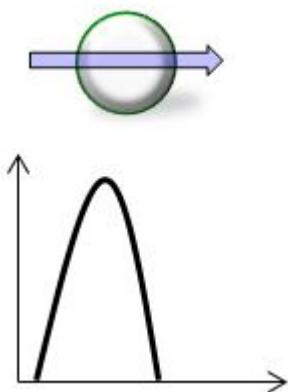
- Finds factor of difference between transcripts not the net difference

Highly Variable Genes

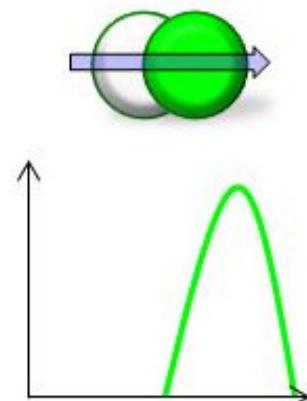
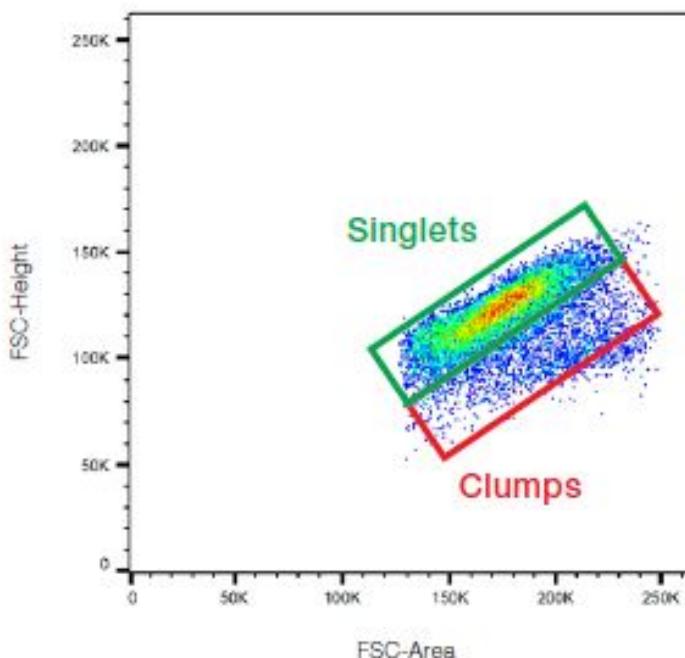
-Selects for the genes highly variable among cells (removes housekeeping genes)



Doublet Filtering



Singlet



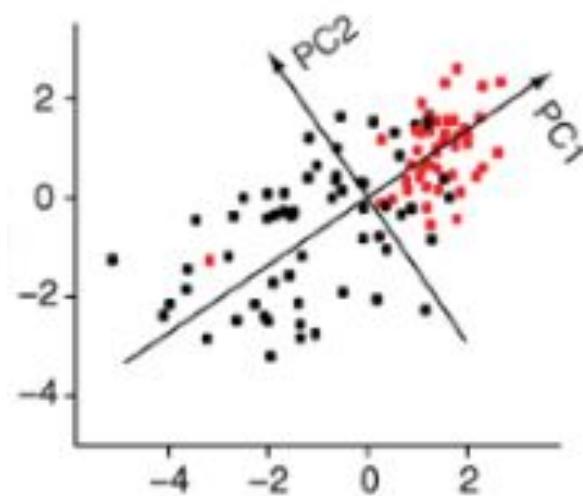
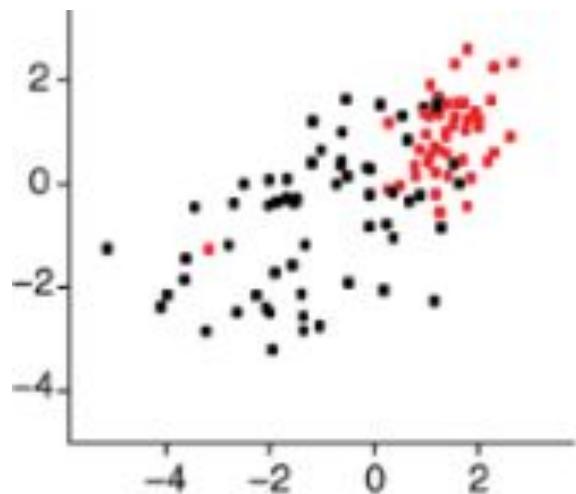
Doublet

scRNA-Seq: Methods

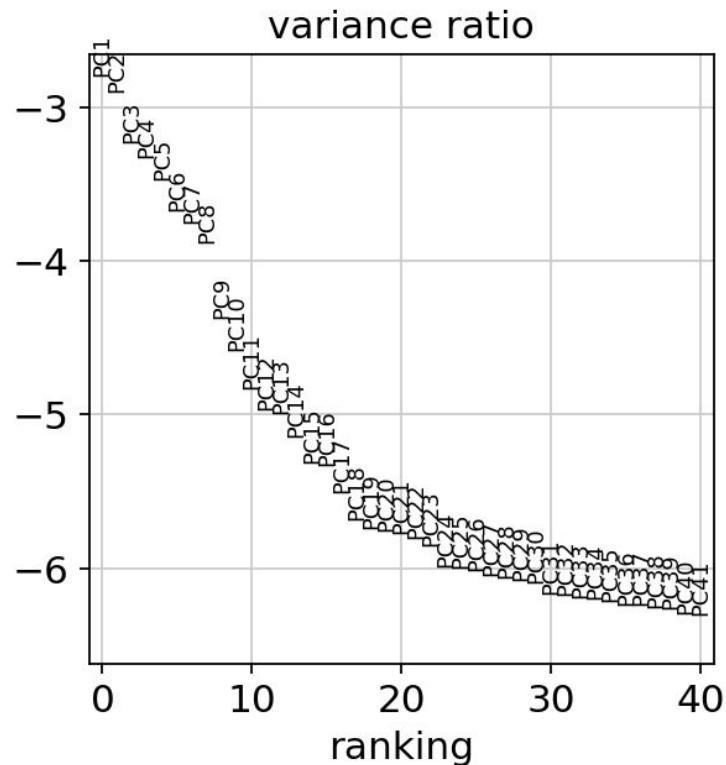
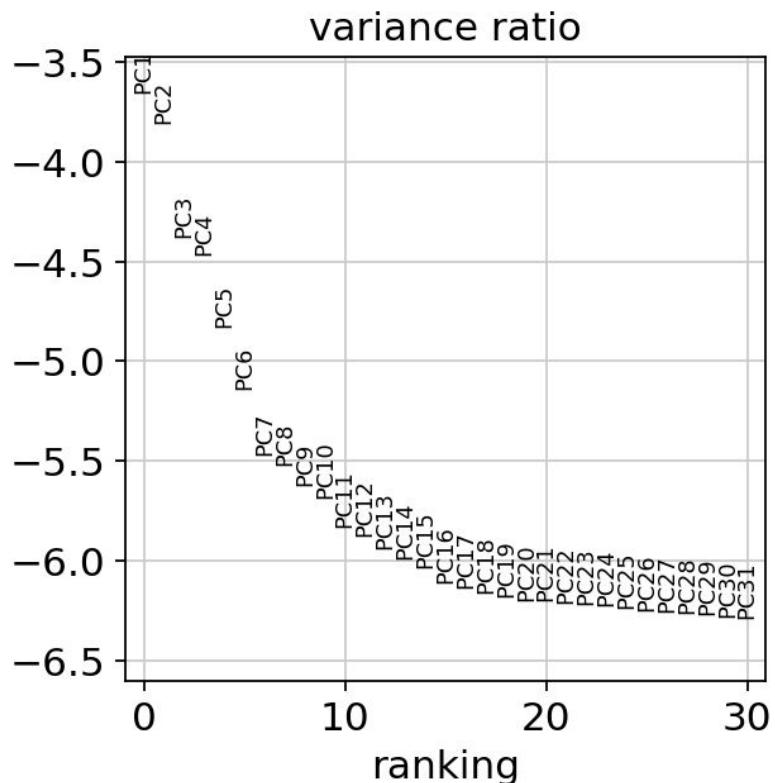
PCA

Principal Component Analysis: PCA

- Reduces the dimensionality of the data
- Allows for finding the components of the data that captures the most variation in the data set while removing technical noise



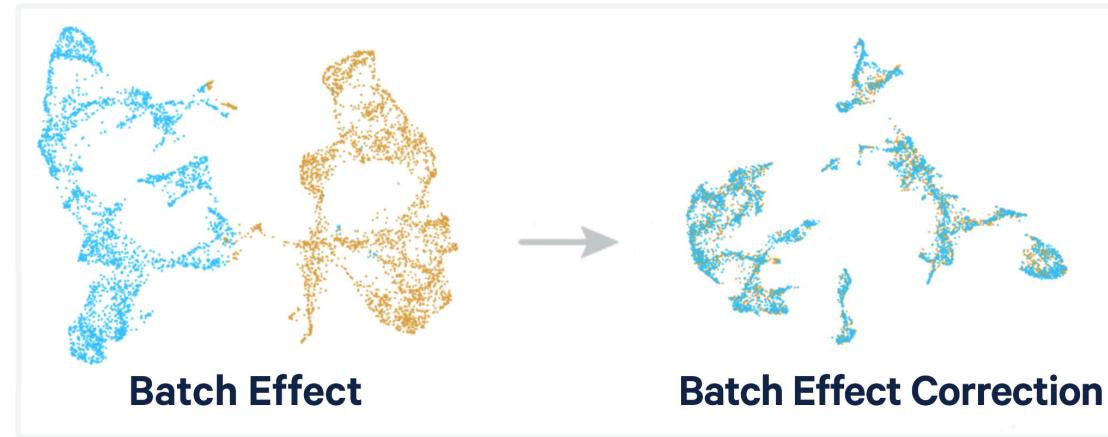
Principal Component Selection



scRNA-Seq: Methods Batch Correction/Clustering

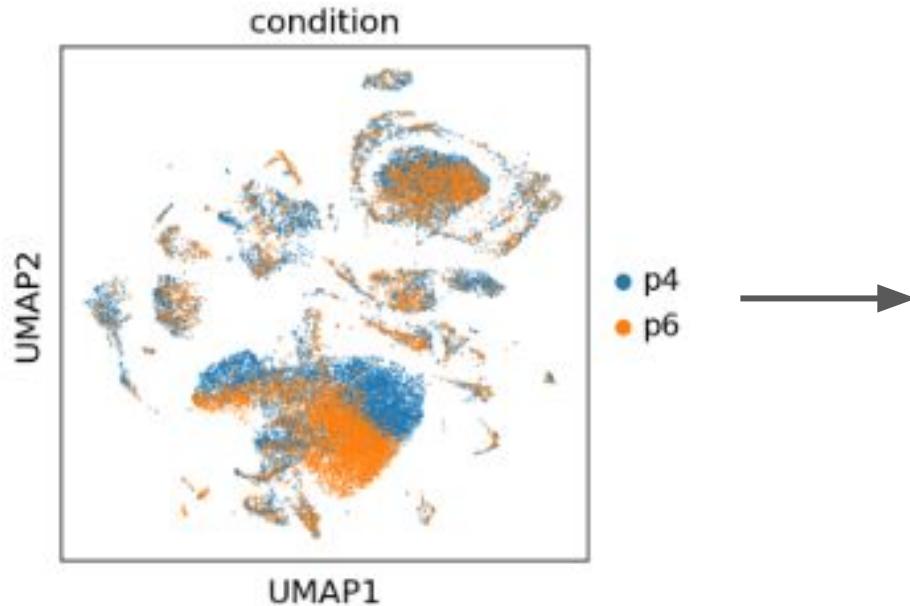
Batch Effect Correction

- Each timepoint except p0 contains two batches (batch #1 and batch #2)
- Batches will have natural variation in expression of genes, transcripts, cell types, etc.
- Integrating batches without accounting for this natural variation can lead to unwanted cluster segregation and higher redundant cluster counts

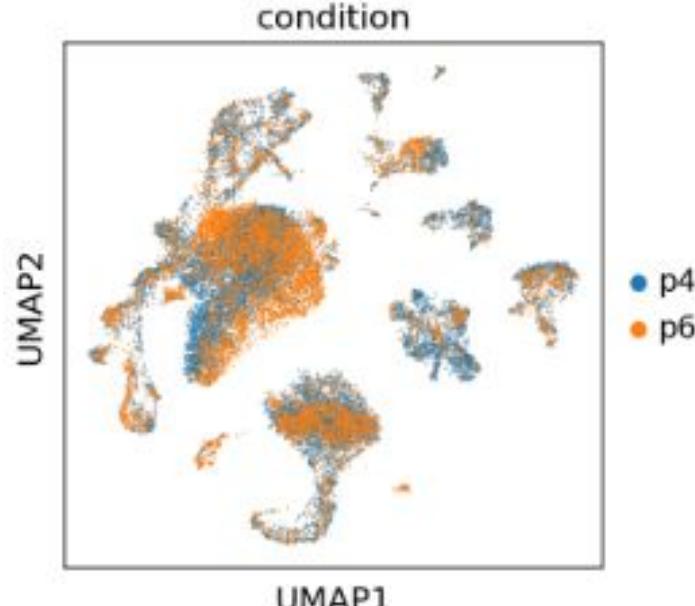


BBKNN Batch Correction

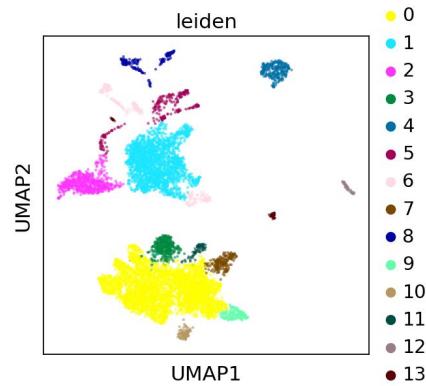
Non-Batch Corrected (p4/6)



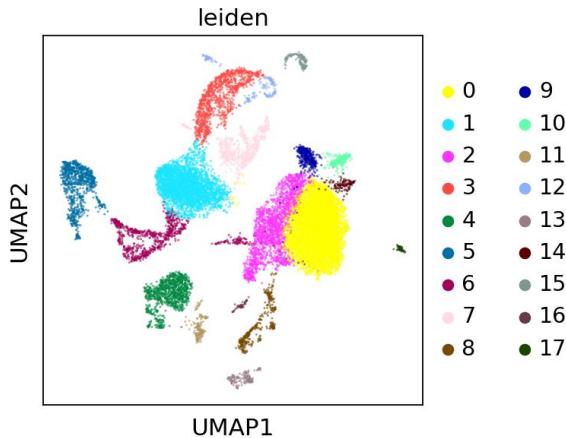
Batch Corrected (p4/6)



p0

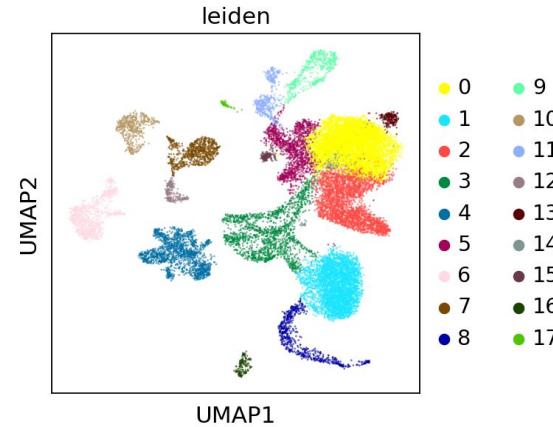


p10

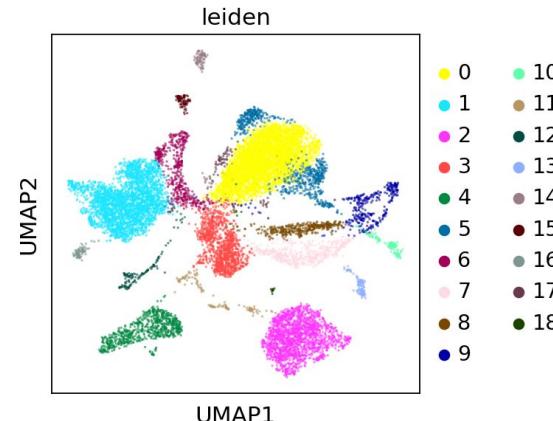


All Timepoints and Cell Types

p4/6



p15



Identifying Neuronal Cells

Finding the Neuronal Cells (Markers)

Syn1

- Regulates neurotransmitter release associated with the small synaptic vesicles

Rbfox3

- Produces the neuronal nuclei (NeuN) antigen used as a marker for post-mitotic neurons

Gad1

- Glutamic acid decarboxylase, responsible for producing the enzyme that catalyzes the production of GABA from L-glutamic acid (Inhibitory Marker)

Slc17a6

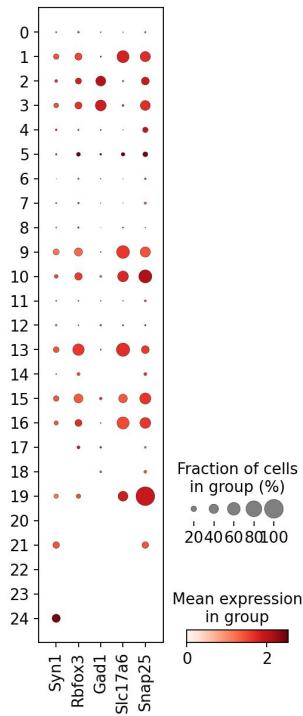
- Enable L-glutamate transmembrane transporter activity and neurotransmitter transmembrane transporter activity (Excitatory Marker)

Snap25

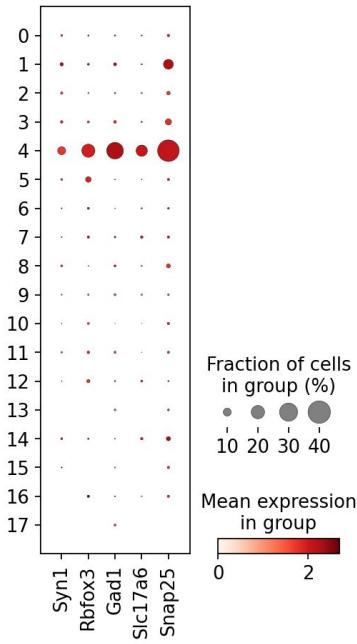
- Plays a role in Synaptogenesis negatively modulates neuronal voltage-gated calcium channels

Identifying Neuronal Cells: Mean Expression Dotplot

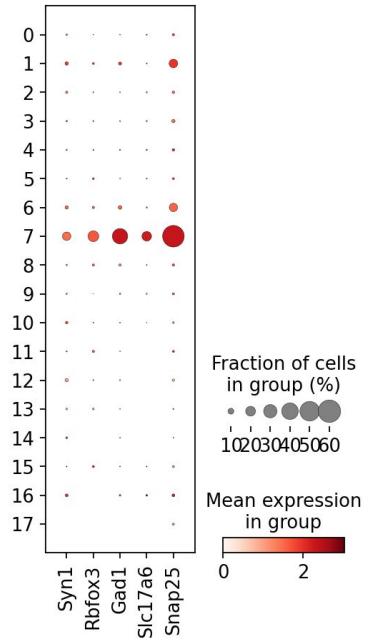
p0
neuronal Genes
Grouped by Cluster



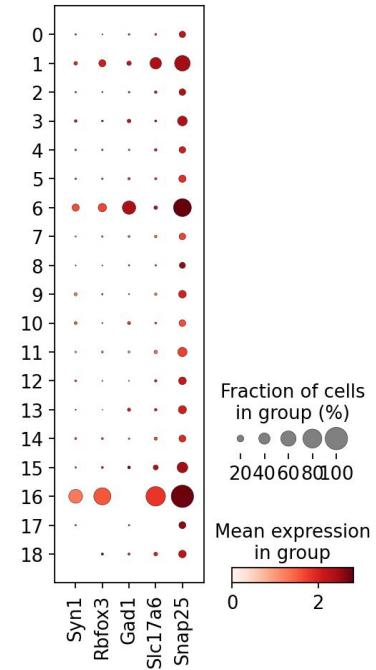
p4-6
neuronal Genes
Grouped by Cluster



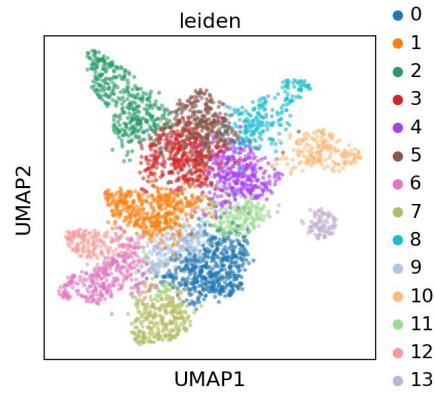
p10
neuronal Genes
Grouped by Cluster



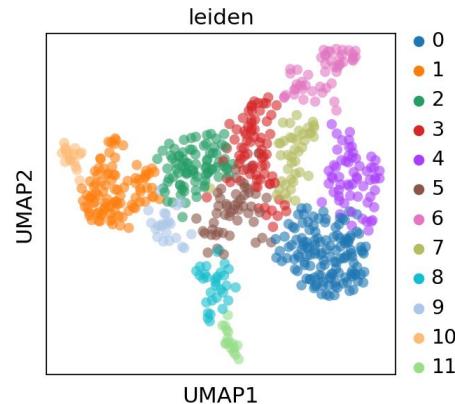
p15
neuronal Genes
Grouped by Cluster



p0

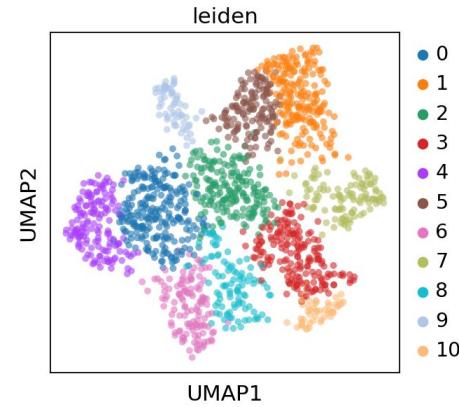


p10

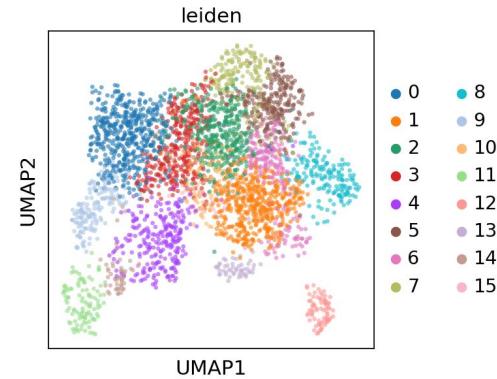


All Timepoints Neuronal Cells

p4/6

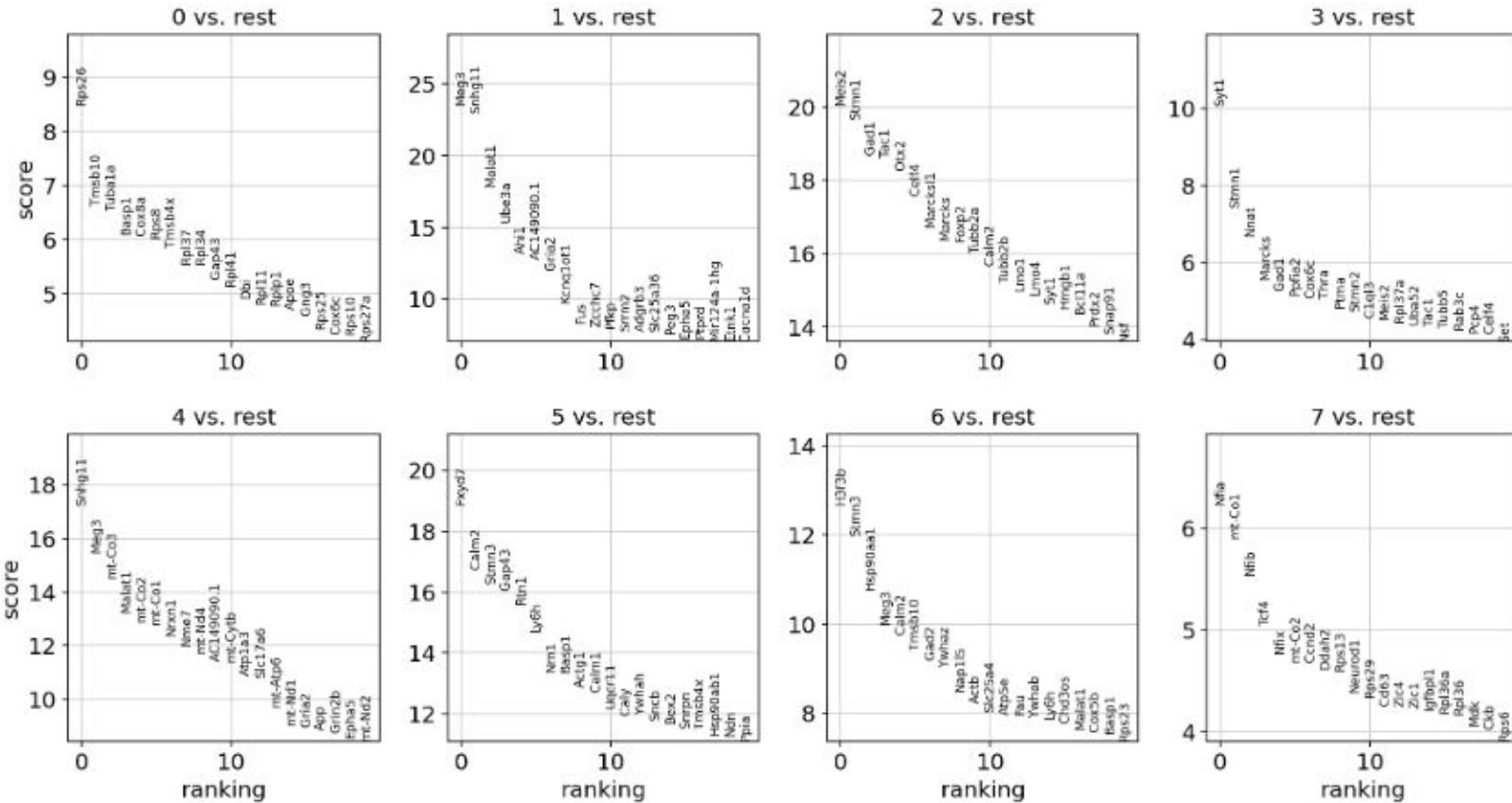


p15



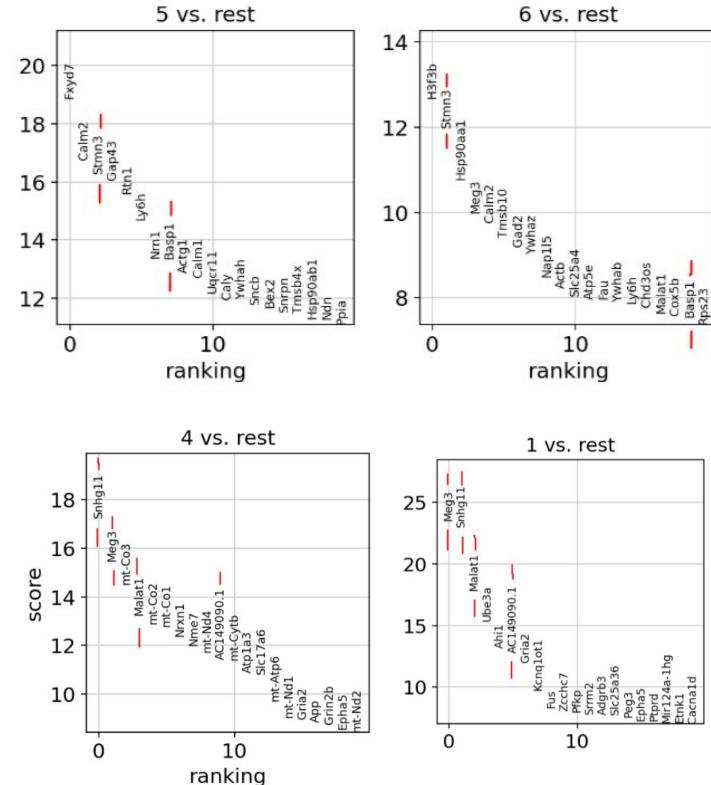
Unique Genes

Rank Genes: Differentially Expressed Genes



Unique Genes for a Neuronal Cluster

- One may think identifying unique genes is as simple as choosing the highest differentially expressed genes for a cluster ($p > .01$)
- Highly differentially expressed genes can be shared between clusters
- A smaller sub-data set means more similar genes



Unique Genes: Ratio Metric Mean Expression Ranking

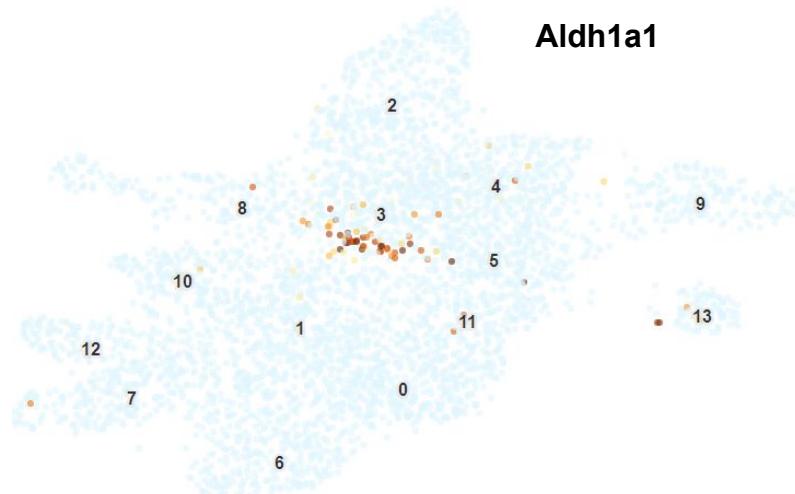
- Extract highly differentially expressed genes ($p < .05$) for each cluster
- For each gene find the mean expression for the cluster in question (ME_C), and the average expression level of the gene, for all other clusters (ME_A)
- $(ME_C) / (ME_A) = \text{ratio of expression of a gene vs all other clusters}$
- Repeat process for all genes for all clusters, and rank the genes from highest ratio to lowest

The top genes will be uniquely expressed within the cluster

p0 TimeFrame Ratio Ranking of Unique Genes

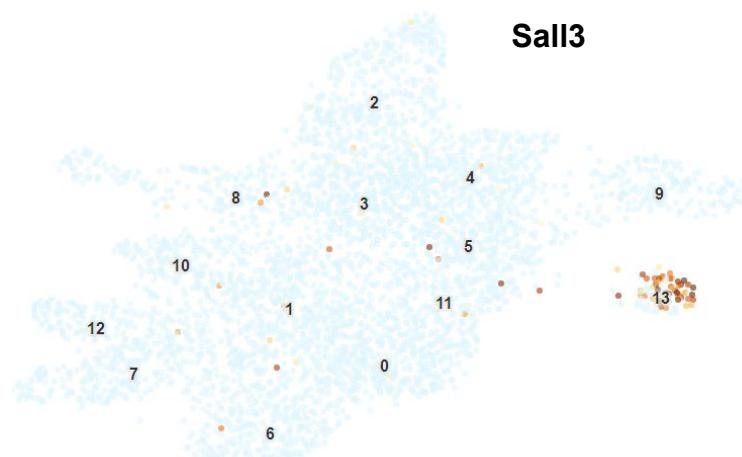
Cluster Number: 3

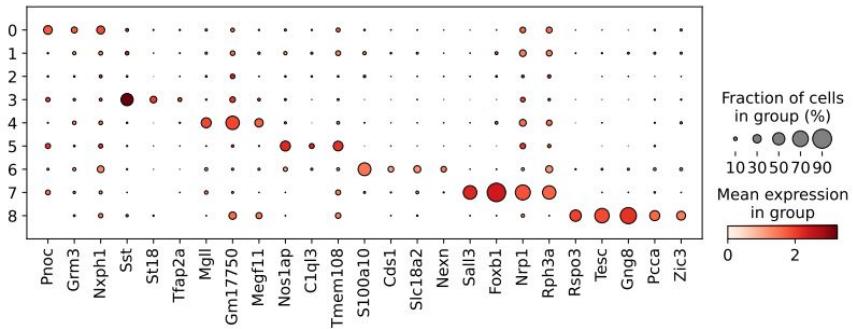
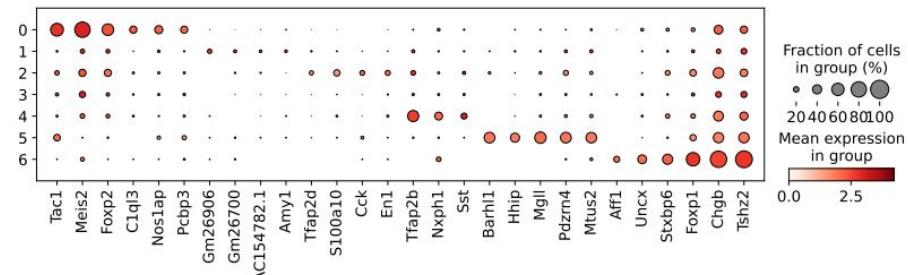
Aldh1a1	106.01914826606604
Ddc	85.74116272852227
Cnpy1	36.82405434722067
Gch1	33.83131950080503
Ret	33.20699886191888
Iqcj	27.823929325074637
En2	18.799957976090347
S1c10a4	18.649614426077047
Fam210b	17.03059477744272
Ptpru	15.231150403653864



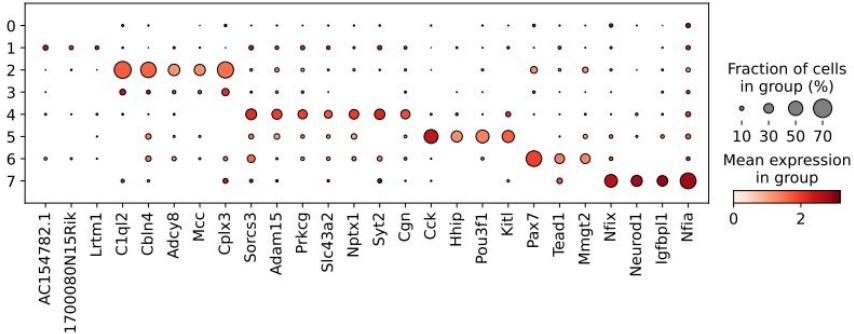
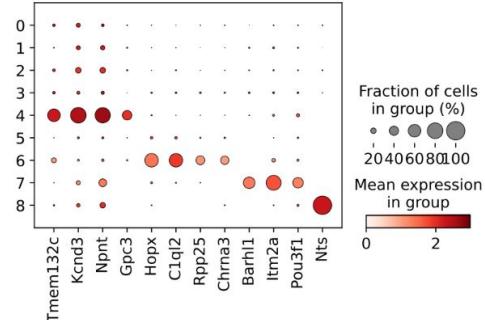
Cluster Number: 13

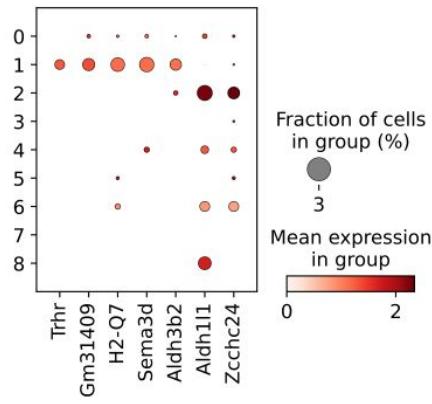
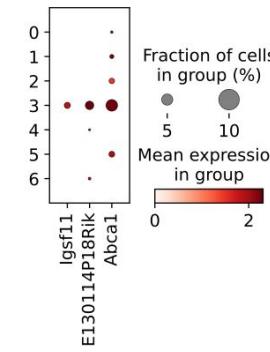
Sall3	105.04409713104727
Tanc1	49.040860483357775
Enpp1	16.012584691898397
B230323A14Rik	11.133931208993701
Mafb	7.4569392497989195
Ryr3	7.158909879796066
Unc5b	6.911838189003894



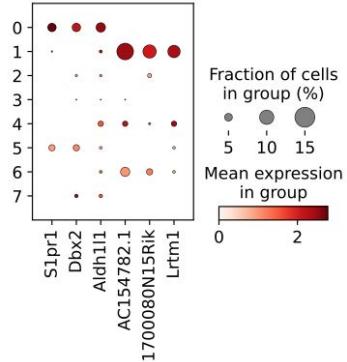
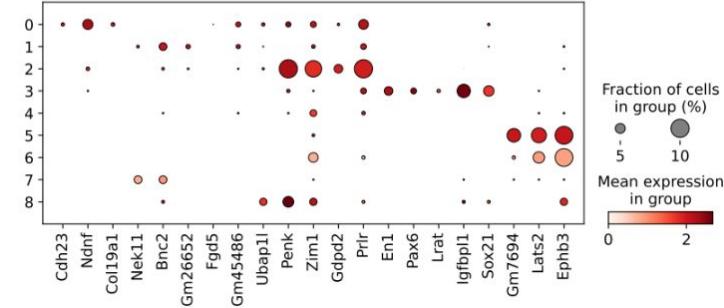
P0**P4/6**

Distinct Unique Genes

P10**P15**

P0**P4/6**

Non-Distinct Unique Genes

P10**P15**

Resolution Adjustment

- Purpose is to minimize the number of non-distinct clusters removing the initial arbitrary count of clusters

Non-Distinct clusters have been defined as having no genes that are....

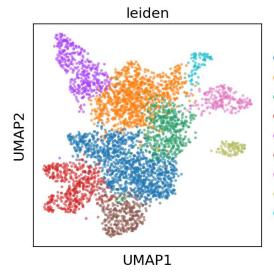
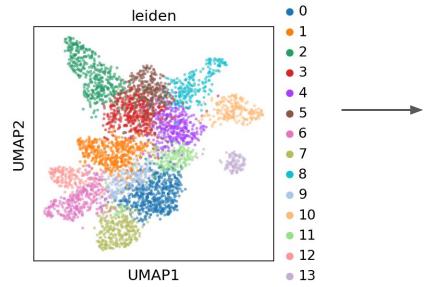
- **Expressed within at least 10% of the cells within given cluster**

AND

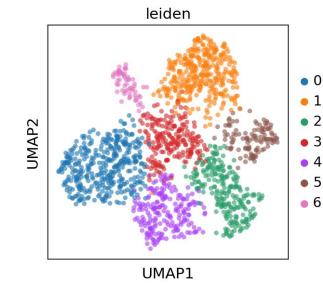
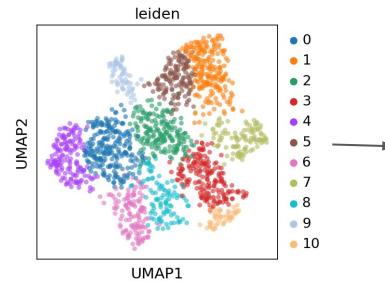
- **Have at least (2X) the percentage of cells than the mean percentage of all other clusters**
- **Have (2X) the mean expression in comparison to all other clusters**

OR

P0

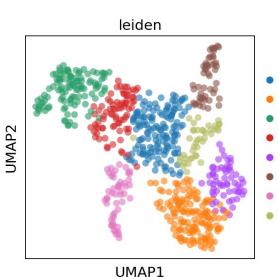
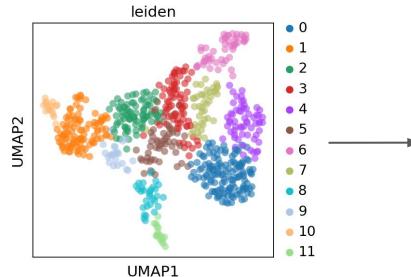


P4/6

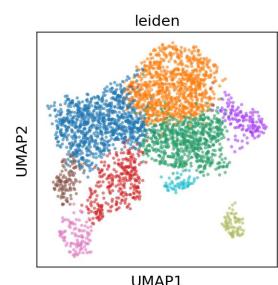
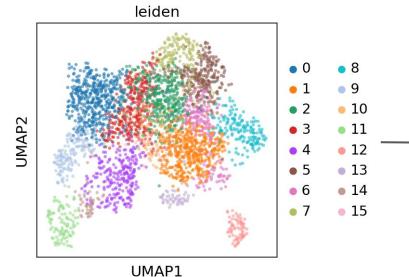


Resolution Adjustment

P10



P15



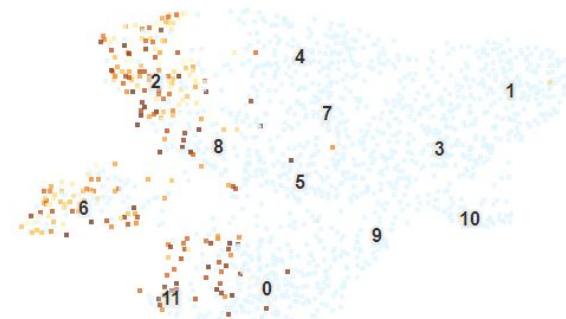
Unique Genes
in some but not
Others

Unique in Some but not Others

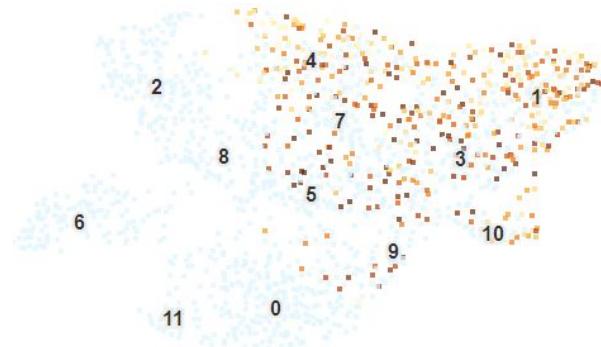
Are there genes expressed highly in more than one cluster that are lowly expressed in the other clusters?

- The neuronal cell groups could easily be divided into excitatory and inhibitory from *Gad1* and *Slc17a6*, is it possible to divide the clusters by other genes markers?

***Slc17a6* (excitatory) p4/6**



***Gad1* (Inhibitory) p4/6**



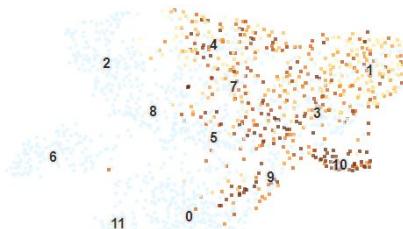
Unique in Some but not Others: Approach

- A simple mean expression ranking does not work because you do not know how many clusters are highly expressing a gene
- As such, you need to quantify the spread of mean expression between groups of clusters.

Standard Deviation Mean Expression Ranking

- Checked whether a highly differentially expressed gene was expressed in more than one cluster
- If true calculated percentage of cells expressing the gene
- Calculated the Standard Deviation of all the percentages, query, then rank genes

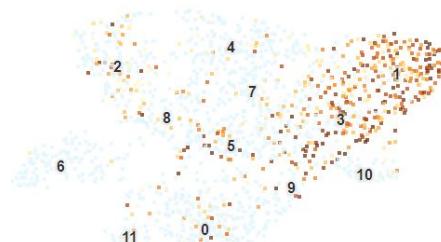
Gata3



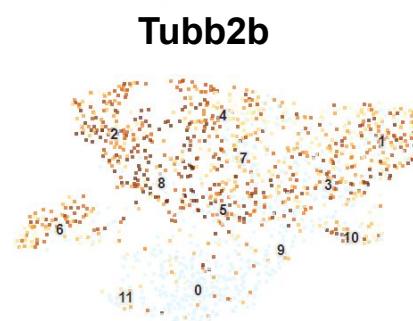
p4/6 TimeFrame Standard Deviation Ranking of Unique Differentially Expressed Genes

Gene:	Gata3	Standard Deviation: 0.34390549388444136
Gene:	Otx2	Standard Deviation: 0.3424787460477524
Gene:	Ebf1	Standard Deviation: 0.31290653327860446
Gene:	Nrn1	Standard Deviation: 0.3029216710586183
Gene:	Nrgn	Standard Deviation: 0.2915018092373296
Gene:	Gap43	Standard Deviation: 0.28778435431789423
Gene:	Syt4	Standard Deviation: 0.27550815628314185
Gene:	Islr2	Standard Deviation: 0.2647449777181341
Gene:	Rorb	Standard Deviation: 0.2624876676049854
Gene:	Lhx1	Standard Deviation: 0.2619163024860561
Gene:	Gfral	Standard Deviation: 0.25293238500365994
Gene:	Slc17a6	Standard Deviation: 0.24571851545152118
Gene:	Snhg11	Standard Deviation: 0.2371079788819624
Gene:	Fxyd7	Standard Deviation: 0.23683035710337197
Gene:	Meis2	Standard Deviation: 0.23644473067532673
Gene:	Ubb	Standard Deviation: 0.23552908426564048
Gene:	Tubb2b	Standard Deviation: 0.23360811368112594
Gene:	Lef1	Standard Deviation: 0.2324175195614443
Gene:	Gad1	Standard Deviation: 0.23148173480279352
Gene:	Btg1	Standard Deviation: 0.23139799296014385
Gene:	Zfhx3	Standard Deviation: 0.23033666539868625
Gene:	Snrspd3	Standard Deviation: 0.22874869019003694
Gene:	Uch11	Standard Deviation: 0.22419617197392797
Gene:	Ptms	Standard Deviation: 0.2236750255217932
Gene:	Bzw1	Standard Deviation: 0.22338259332410973
Gene:	Psmcl	Standard Deviation: 0.22225828330822145

Meis2



Tubb2b



Why are we doing all this?

We don't know how many cell types there are

Some groups of clusters may be a single cell type

Other single clusters may be a unique cell type

(Specific Grouping)

Unique in Some Clusters but not Others

Neuronal Subcell Types

Unique solely to a Cluster

(Broad Grouping)

If a cluster has a poor set of unique genes then the standard deviation ranked genes can help it group with another cluster

If a cluster has a healthy set of unique genes these genes can be searched to catalogue a cell type

Allen Brain Atlas

Allen Brain Atlas/Panther/Cell Browser API

p4/6 TimeFrame Ratio Ranking of Unique Differentially Expressed Genes

Cluster Number: 2

Pitx2:	9.000000000000018
Slc38a4:	8.000000000000016
Prss12:	7.000000000000007
Iqcf3:	5.000000000000001
Efcab10:	5.000000000000001
Wif1:	4.500000000000009
Neurod6:	4.499999999999992
Gal:	4.000000000000006
Pou4f2:	4.000000000000004
Meig1:	4.000000000000004
1700016K19Rik:	3.500000000000001
Nkxp4:	3.3333333333327
Sncg:	3.000000000000004
Dlk1:	3.000000000000004
Fibin:	3.000000000000036
Nhlh2:	2.624999999999982
Myh14:	2.600000000000005

p4/6 TimeFrame Standard Deviation Ranking of Unique Differentially Expressed Genes

Gene: Gata3	Standard Deviation: 0.34416826446430243
Gene: Ctx3	Standard Deviation: 0.3428860972679531
Gene: Ebf1	Standard Deviations: 0.31179931326756216
Gene: Nrn1	Standard Deviations: 0.30379410512212174
Gene: Gap43	Standard Deviations: 0.28908094020305663
Gene: Nrgn	Standard Deviations: 0.2871773870302733
Gene: Syt4	Standard Deviations: 0.27462618673991196
Gene: Lhx1	Standard Deviations: 0.26217565011158727
Gene: Islr2	Standard Deviations: 0.26146310127161009
Gene: Rorb	Standard Deviations: 0.2602956024978234
Gene: Grf1	Standard Deviations: 0.25411503123074075
Gene: Slc17a6	Standard Deviations: 0.25114568591494313
Gene: Ubb	Standard Deviations: 0.24046991218944327
Gene: Shh11	Standard Deviations: 0.2389168642299025
Gene: Fwyd7	Standard Deviations: 0.2378575162126161908
Gene: Meis2	Standard Deviations: 0.2368512030102427
Gene: Gad1	Standard Deviations: 0.23217570056777057
Gene: Igf1	Standard Deviations: 0.23215769257277332
Gene: Lef1	Standard Deviations: 0.23177058155165046
Gene: Zfhx3	Standard Deviations: 0.23128146592215052
Gene: Tubb2b	Standard Deviations: 0.23119468677438892
Gene: Snarp3	Standard Deviations: 0.23093340951922026

Unique

SD



Gene

Brain Section Image



Otx2

Cell Browser Exp. Profile



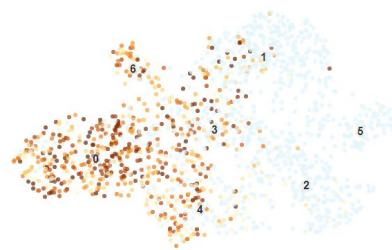
Ontologies

Molecular/Biological/Cellular Ontology Terms:

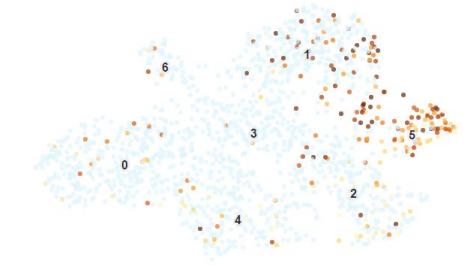
- DNA-binding transcription factor activity, RNA polymerase II-specific GO:00000081
- RNA polymerase II transcription regulatory region sequence-specific DNA binding GO:0000077
- regulation of transcription by RNA polymerase II GO:00000357
- transcription by RNA polymerase II GO:00000166
- membrane GO:0005576
- RNA polymerase II transcription regulatory region sequence-specific DNA binding GO:0000077
- DNA-binding transcription factor activity, RNA polymerase II-specific GO:00000801
- RNA polymerase II transcription regulatory region sequence-specific DNA binding activity, RNA polymerase II-specific GO:0001228
- DNA binding GO:00000677
- sequence-specific double-stranded DNA binding GO:00000764
- regulation of DNA-templated transcription GO:00000357
- positive regulation of transcription by RNA polymerase II GO:00000357
- sensory perception of sound GO:0007695
- midbrain development GO:0003691
- regulation of process of cell differentiation GO:0000524
- positive regulation of transcription by RNA polymerase II GO:0001340
- neuron apoptotic process GO:1396584
- negative regulation of outer hair cell apoptotic process GO:1980586
- nucleus GO:0000334

P4: Genes Concentrated In Superficial SC

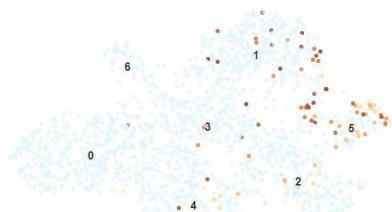
Otx2



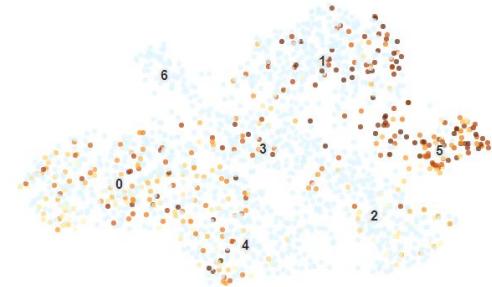
Gfra1



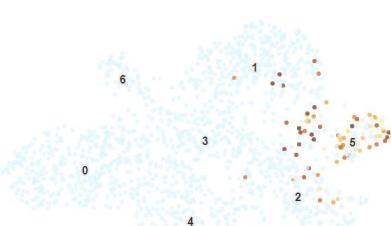
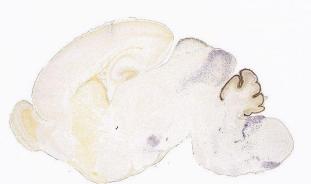
Cdh7



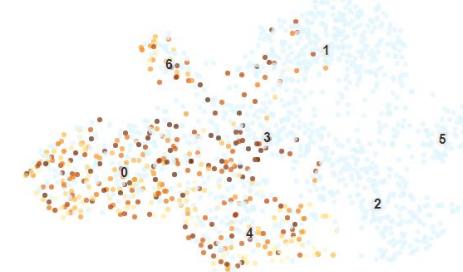
Rorb



Barhl1



Gad1

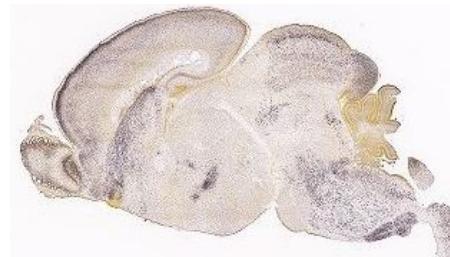


P4: Genes Concentrated in the Intermediate/Deep SC

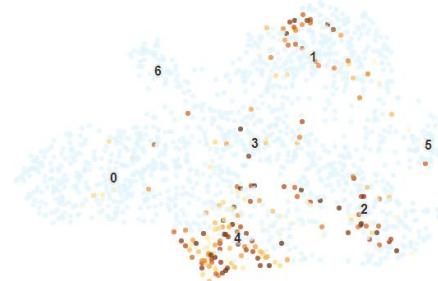
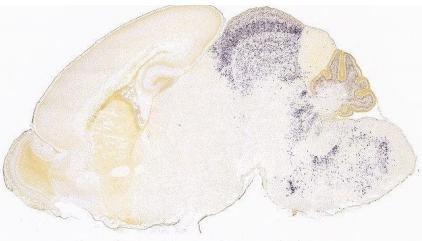
Ebf3



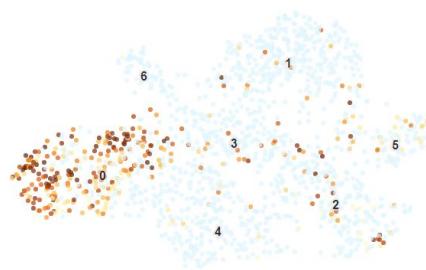
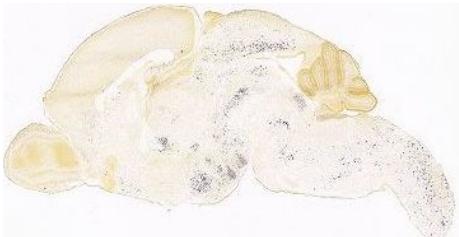
Meis2



Tfap2b



Tac1



P4: Clusters Concentrated In the SC

- Cluster 0, 4, 6, and 5 are all possible candidates with being primarily present in the Superficial SC.
- Cluster 0 and 5 are the strongest.
- The Intermediate and Deep layers have possible candidates of Cluster 4, Cluster 6, and Cluster 0

P4 BEST SECTIONS			
Gene	Heavy Cluster(s)	Light Clusters(s)	SC Concentration
Otx2	0,6,3,4,	1,	Superficial
Gad1	0,4,6	1,3	Superficial
Cdh7	5,	2,1,	Superficial
Barhl1	5,	2,1,	Superficial
Gfra1	5,1	2,0,4	Superficial
Rorb	5,1,0,	2,3,4	Superficial
Meis2	0,1,2,3	4,	Intermediate/Deep
Ebf3	6,5	2,1,3	Intermediate/Deep
Tfap2b		4,2,1	Intermediate/Deep
Tac1		0,2,3,5	Deep?

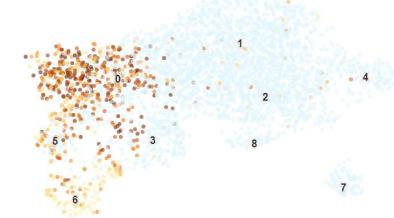
P4: Other Clusters Concentrated In the SC

Is cluster 2 possibly correlated more with the inferior colliculus?

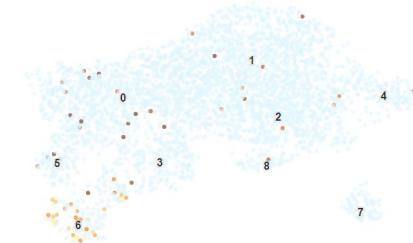
P4 OKAY SECTIONS			
Gene	Heavy Cluster(s)	Light Clusters(s)	SC Concentration
Cntn3		5,1,2	Superficial
Gata3	6,0,4,	1,3	Superficial (Light everywhere)
Hhip	5,	1,	Superficial
Nnat	0,2,3,4,5,6,	1,	Superficial (Mod everywhere)
Sox4	0,4,3,6,2,	1,5,	Superficial (Mod everywhere)
Tubb3	0,2,4,5,6	1,	Superficial (Mod everywhere)
Barhl2	None	2,1	Superficial/Intermediate (Mod everywhere)
Gap43	0,1,2,3,4,5		Superficial/Intermediate (Mod everywhere)
Marcks1	0,2,3,4,5,6,	1,	Superficial/Intermediate (Mod everywhere)
Foxb1	None	2,1,3	Intermediate
Dkk3	5,	1,2,4	Intermediate/Deep (Mod everywhere)
Slc17a6	5,2	1,	Deep/Intermediate
Htr2c	None	6,1,2,4	Deep
Zfhx3	0,4,2,1	3,	Superficial/Deep
Tfap2d	2,	1,	Inferior Colliculus
En1	2,	4,	Inferior Colliculus

P14: Genes Concentrated in the Superficial SC

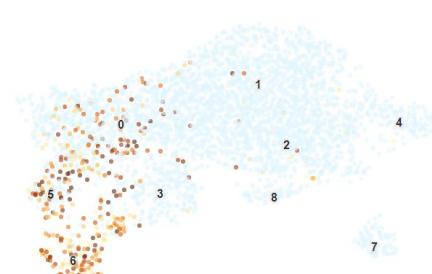
Slc32a1



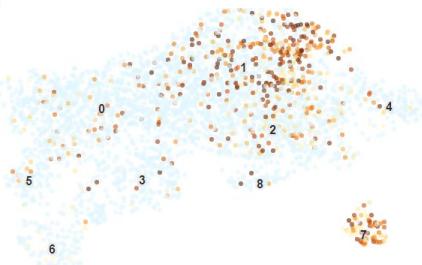
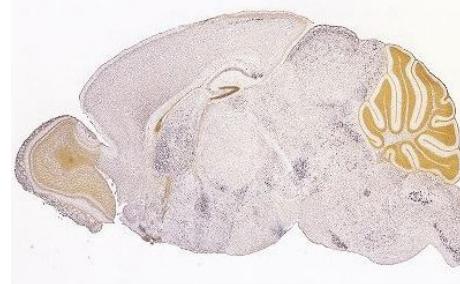
Chrna3



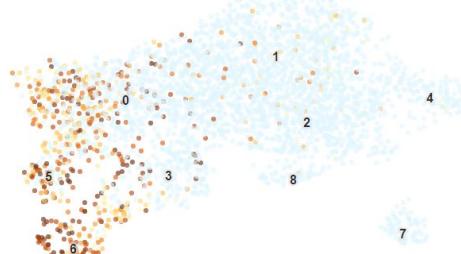
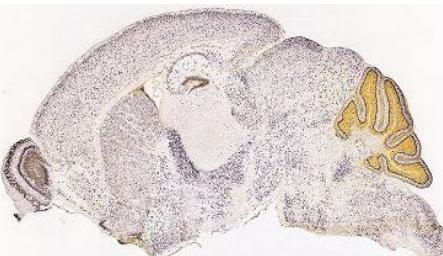
Otx2



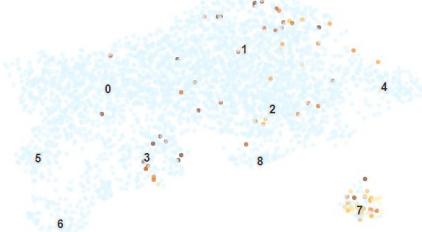
Gfra1



Gad1

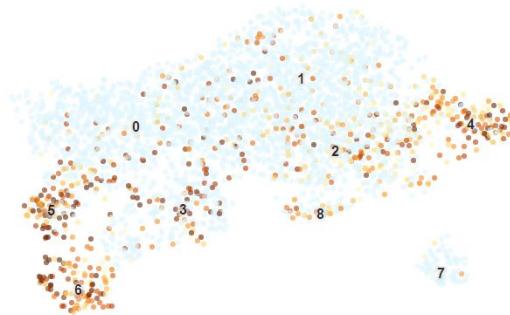
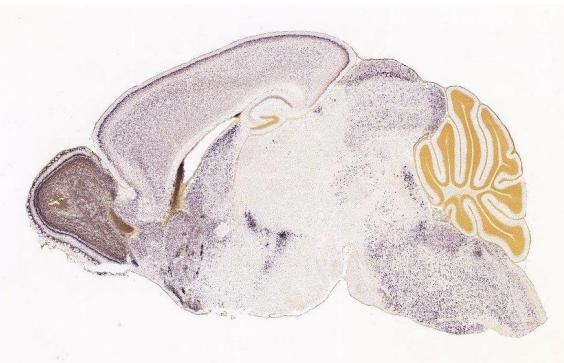


Barhl1

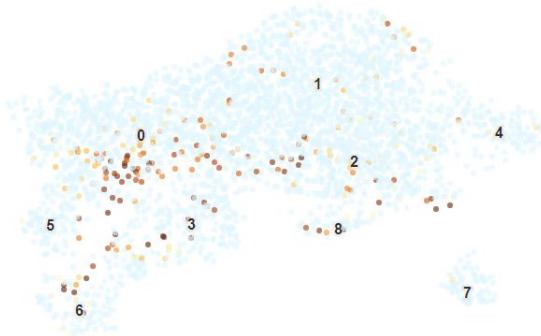
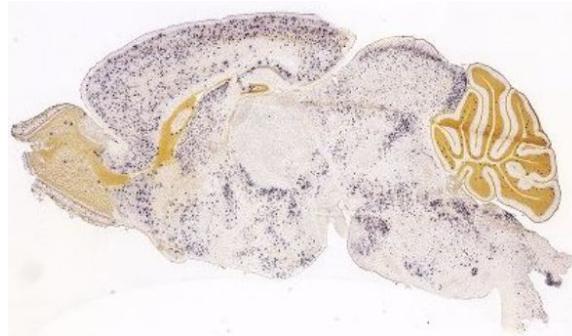


P14: Genes Concentrated in the Intermediate/Deep SC

Meis2



Sst



P14: Clusters Concentrated In the SC

- Cluster 0, 7, 6, and 5 are all possible candidates with being primarily present in the Superficial SC.
- The Intermediate and Deep layers have possible candidates of Cluster 5, Cluster 6, Cluster 4, and Cluster 0

P14 BEST SECTIONS			
Gene	Heavy Cluster(s)	Light Clusters(s)	SC Concentration
Otx2	6,5,0		Superficial
Slc32a1	5,6,0	3,	Superficial/Intermediate (Mod Everywhere)
Gad1	0,5,6	1,3	Superficial
Tac1	7,6,5	1,2	Superficial
Barhl1	7,		Superficial
Gfra1	7,1	2,0,4	Superficial
Chma3	6,	0,	Superficial/Intermediate
Meis2	5,6,4	2,1,0	Intermediate/Deep
Sst	0,8	3,6,2	Intermediate/Deep

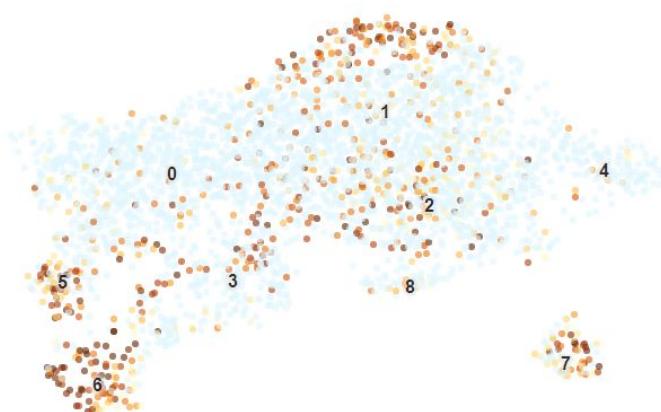
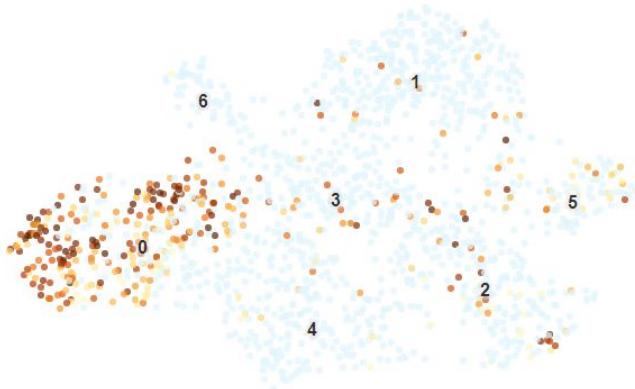
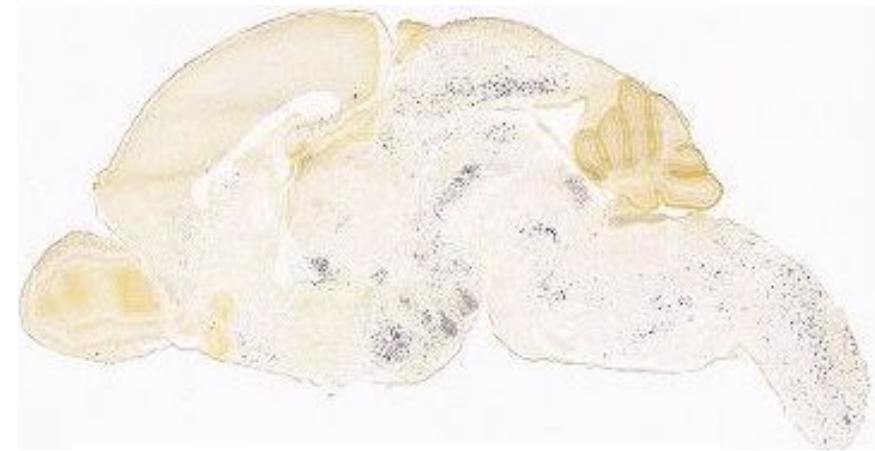
P14: Other Clusters Concentrated In the SC

P14 OKAY SECTIONS			
Gene	Heavy Cluster(s)	Light Clusters(s)	SC Concentration
C1q12	6,	5,3	Superficial
Gabbr1	6,	5,	Superficial
Gad2	5,6,0	3,	Superficial/Intermediate (Mod Everywhere)
Kcnab1	7,	1,5,0,2	Superficial/Intermediate
Marcks1	6,7,3,5	0,1,2,4	Superficial/Intermediate (Mod Everywhere)
Sema3e	4,	0,2,1	Intermediate
Sic17a6	1,2,4,8,7	3,	Intermediate/Deep (Mod Everywhere)

Tac1: P4 vs P14

P4

P14

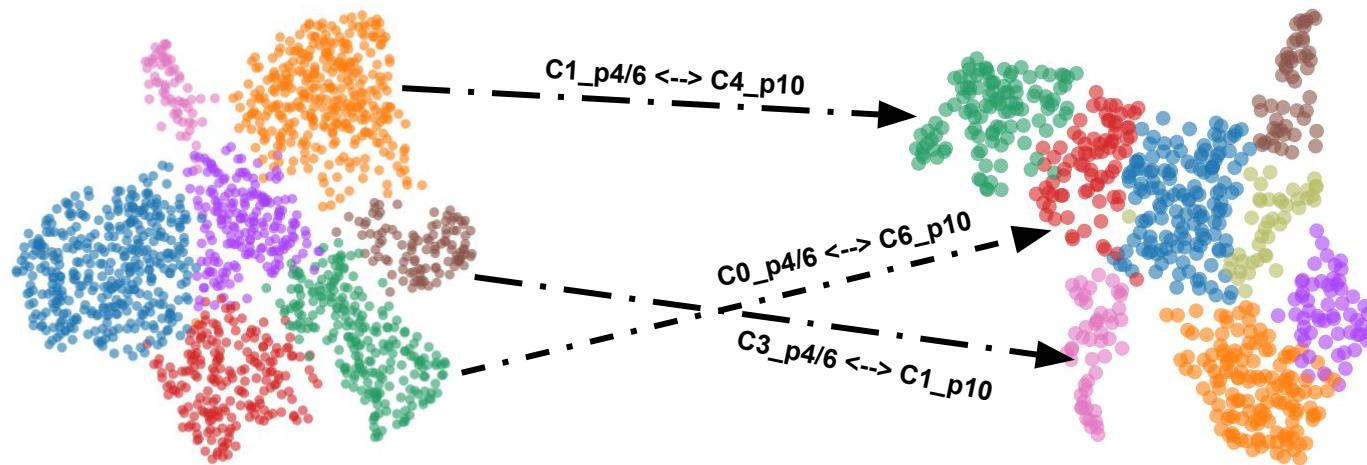


Cluster Relationships through Time

Finding Cluster Relationships through Time

Can we match clusters from one timepoint to a cluster of another timepoint?

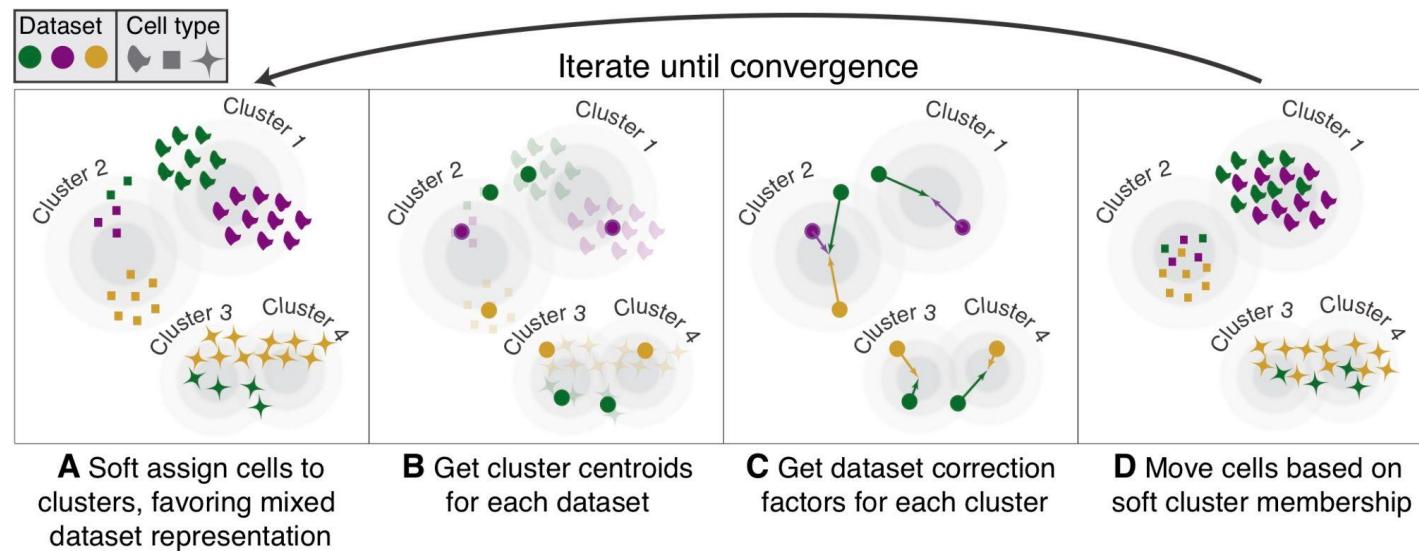
In other words can we determine a cluster's closest clusteral neighbor from another timepoint?

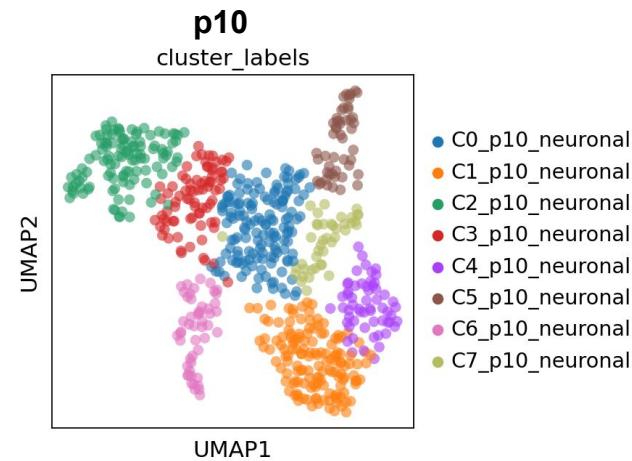
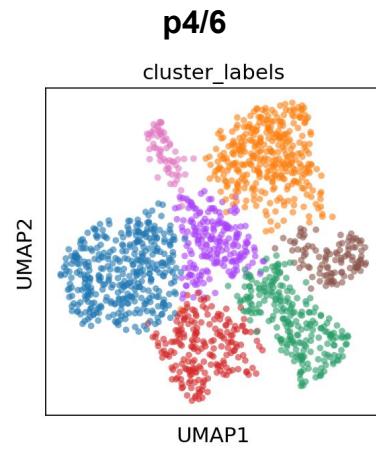


Finding Cluster Relationships by Integrating Timepoints

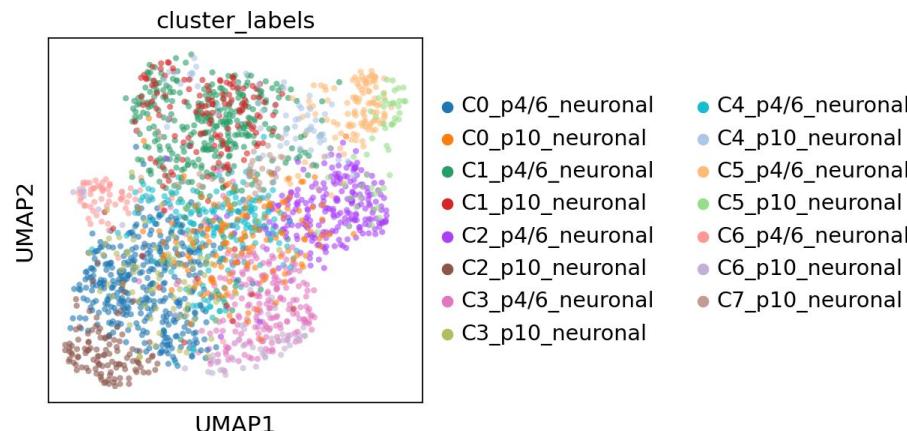
- Better method for quantifying relationships between clusters and timepoints and well as excluding the timepoint in question

Timepoint Integration through Harmony

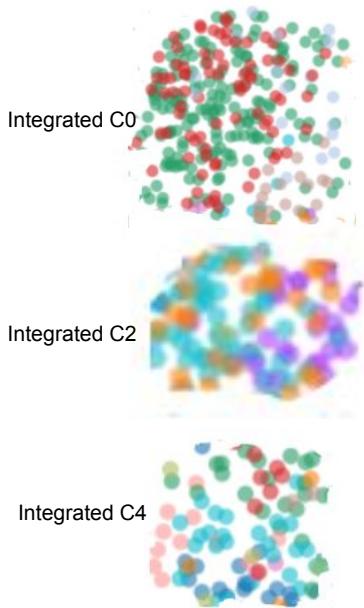




p4/6 vs p10 Integrated Dataset



Integrate Timepoints: Approach



C1_p4/6 2.3% <--> C2_p4/6 18.3% weight: 20.6
C1_p4/6 2.3% <--> C4_p10 28.3% weight: 30.6
C1_p4/6 2.3% <--> C8_p4/6 78.3% weight: 80.6

C1_p4/6 54.1% <--> C7_p10 8.2% weight: 62.3
C1_p4/6 54.1% <--> C4_p4/6 88.1% weight: 142.2
C1_p4/6 54.1% <--> C6_p10 78.9% weight: 133

C1_p4/6 36.7% <--> C2_p10 2.3% weight: 39
C1_p4/6 36.7% <--> C4_p10 21.5% weight: 58.2
C1_p4/6 36.7% <--> C8_p4/6 79.9% weight: 116.6

Cluster label in Question:
C1_p4/6

Repeat for next cluster label

- Strip out clusters from same timepoint
- Order and Rank highest percentage pairs

C1_p4/6 54.1% <--> C6_p10 78.9% weight: 133
C1_p4/6 54.1% <--> C7_p10 8.2% weight: 62.3
C1_p4/6 36.7% <--> C4_p10 21.5% weight: 58.2
C1_p4/6 36.7% <--> C2_p10 2.3% weight: 39
C1_p4/6 2.3% <--> C4_p10 28.3% weight: 30.6

p4/6 vs p10 Clusteral Relationship (>10% of cells)

p10 Cluster 0:

```
C0_p10_neuronal: 60.96% <----> C4_p4/6_neuronal: 61.50%
C0_p10_neuronal: 18.49% <----> C3_p4/6_neuronal: 86.27%
C0_p10_neuronal: 60.96% <----> C2_p4/6_neuronal: 28.99%
C0_p10_neuronal: 10.96% <----> C0_p4/6_neuronal: 69.01%
C0_p10_neuronal: 60.96% <----> C0_p4/6_neuronal: 13.02%
```

p4/6 Cluster 0:

```
C0_p4/6_neuronal: 69.01% <----> C2_p10_neuronal: 100.00%
C0_p4/6_neuronal: 69.01% <----> C3_p10_neuronal: 50.63%
C0_p4/6_neuronal: 11.46% <----> C6_p10_neuronal: 84.00%
C0_p4/6_neuronal: 69.01% <----> C0_p10_neuronal: 10.96%
C0_p4/6_neuronal: 13.02% <----> C7_p10_neuronal: 47.62%
C0_p4/6_neuronal: 13.02% <----> C3_p10_neuronal: 26.58%
C0_p4/6_neuronal: 11.46% <----> C3_p10_neuronal: 18.99%
C0_p4/6_neuronal: 11.46% <----> C0_p10_neuronal: 18.49%
```

p4/6 Cluster 1:

```
C1_p4/6_neuronal: 57.06% <----> C1_p10_neuronal: 64.79%
C1_p4/6_neuronal: 57.06% <----> C4_p10_neuronal: 29.31%
C1_p4/6_neuronal: 21.33% <----> C4_p10_neuronal: 24.14%
C1_p4/6_neuronal: 21.33% <----> C1_p10_neuronal: 22.54%
```

Finished Relational Dataset Output

Correlating Unique Genes Through Time

P4/6

C5_p4/6_neuronal
C1_p4/6_neuronal
C3_p4/6_neuronal
C0_p4/6_neuronal
C4_p4/6_neuronal
C2_p4/6_neuronal
C6_p4/6_neuronal



Timepoint Cluster Pairs

C4_p4/6_neuronal - C0_p10_neuronal
C6_p4/6_neuronal - C5_p10_neuronal
C1_p4/6_neuronal - C0_p10_neuronal
C0_p4/6_neuronal - C3_p10_neuronal
C5_p4/6_neuronal - C3_p10_neuronal
C4_p4/6_neuronal - C4_p10_neuronal
C1_p4/6_neuronal - C4_p10_neuronal
C0_p4/6_neuronal - C2_p10_neuronal
C5_p4/6_neuronal - C6_p10_neuronal
C6_p4/6_neuronal - C6_p10_neuronal
C6_p4/6_neuronal - C2_p10_neuronal
...
...
...

P10

C4_p10_neuronal
C1_p10_neuronal
C0_p10_neuronal
C3_p10_neuronal
C7_p10_neuronal
C2_p10_neuronal
C6_p10_neuronal
C5_p10_neuronal



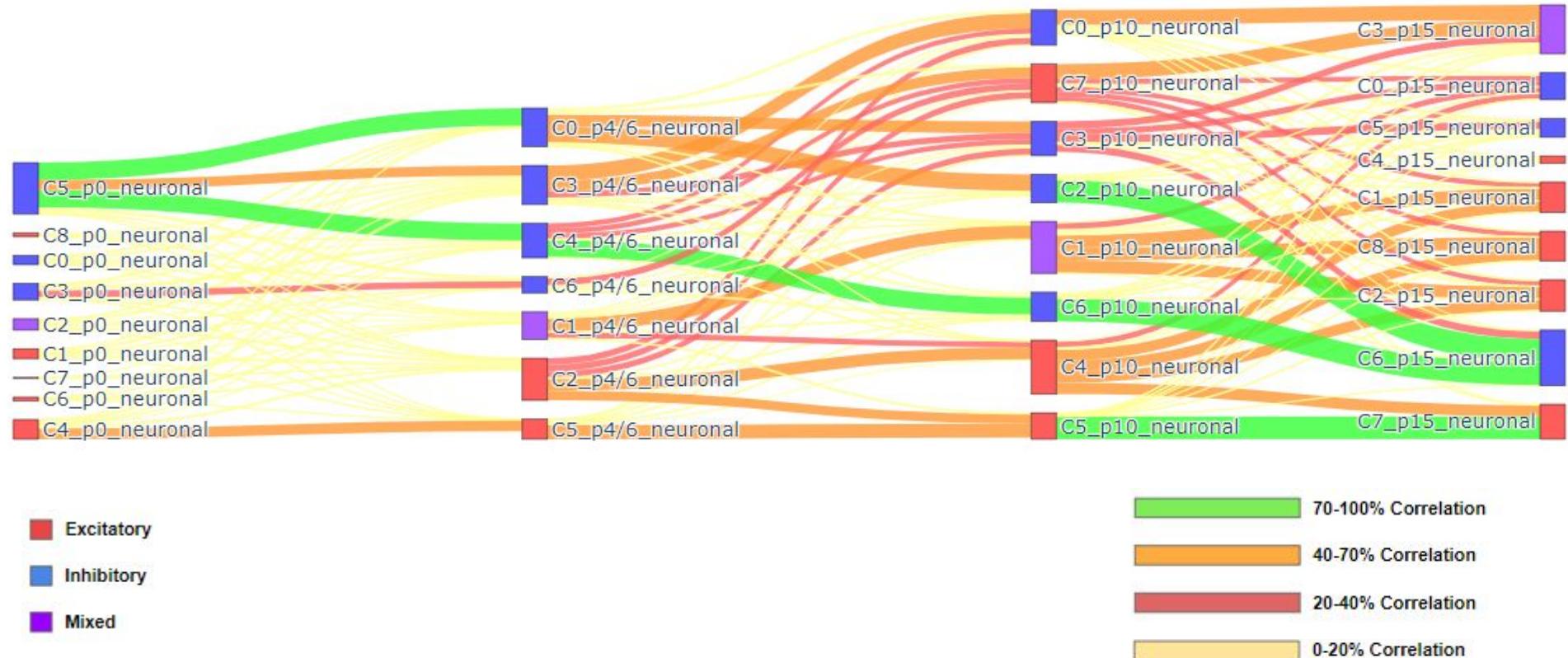
-Find the intersection of all genes within each pair

-Perform Dot-plots

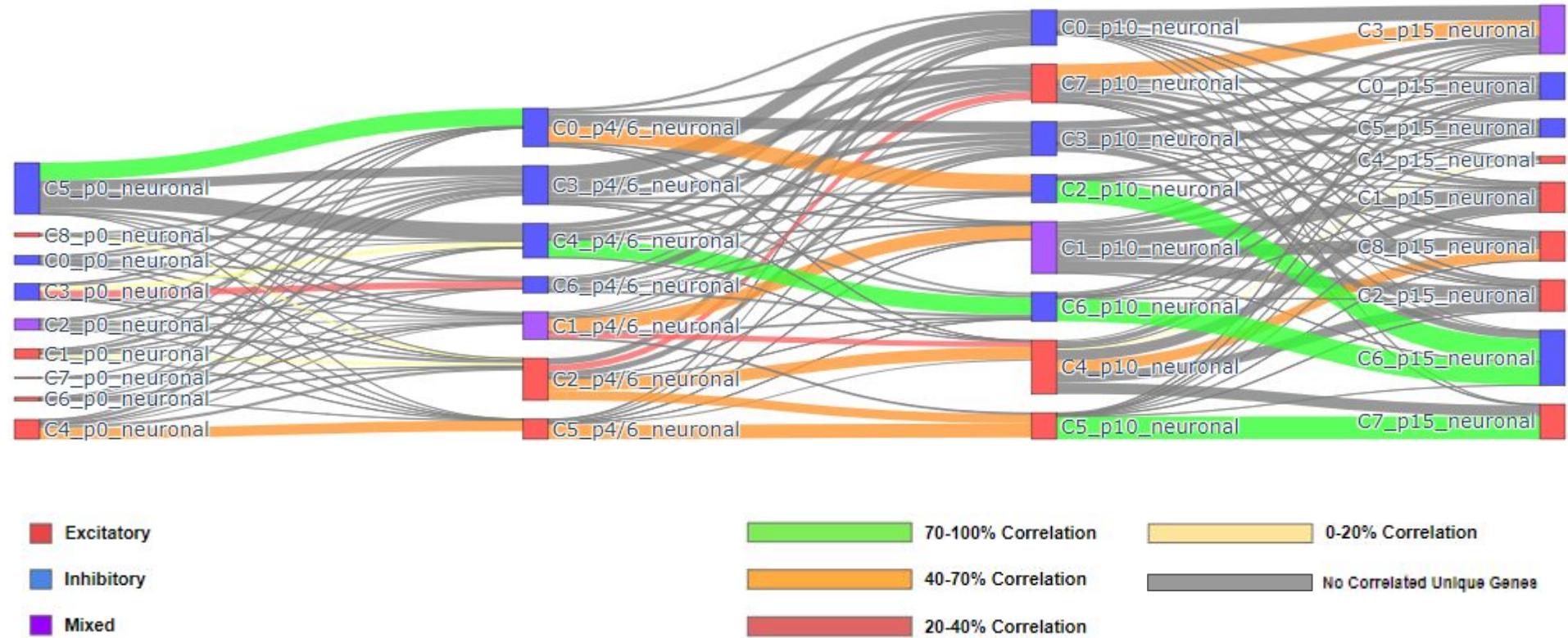
-Remove poor quality genes

-Query genes

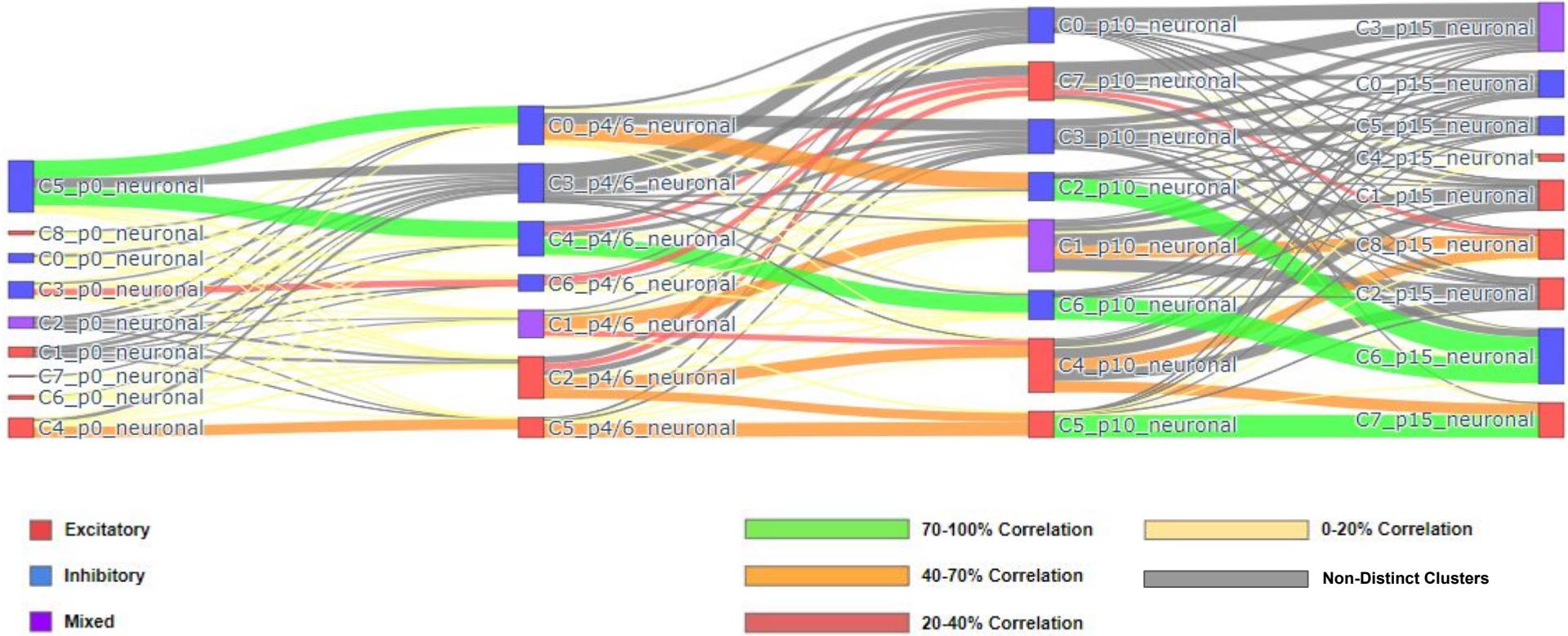
Harmony Integrated Correlations



Harmony Integrated Correlations vs. Correlated Unique Genes



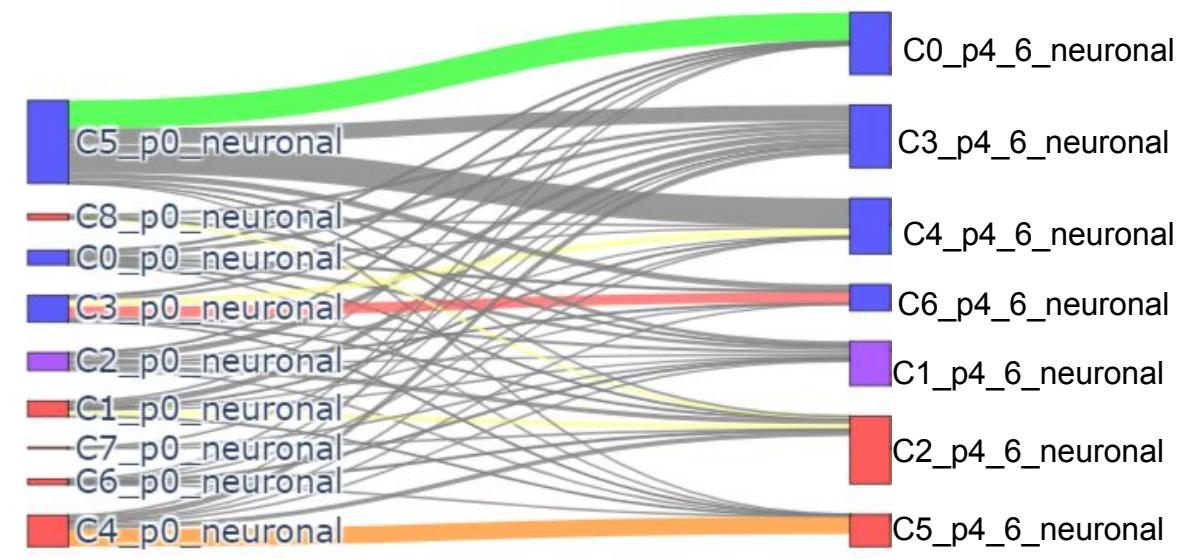
Harmony Integrated Correlations vs. Distinct Clusters



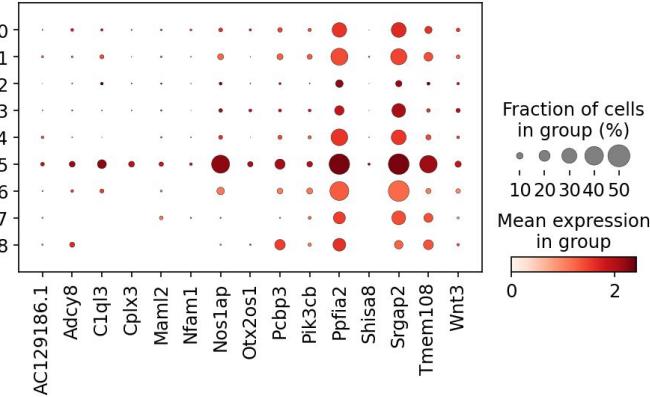
P0 vs P4/6

Correlated Unique Genes

C5_P0 vs C0_P4/6 Correlated Unique Genes

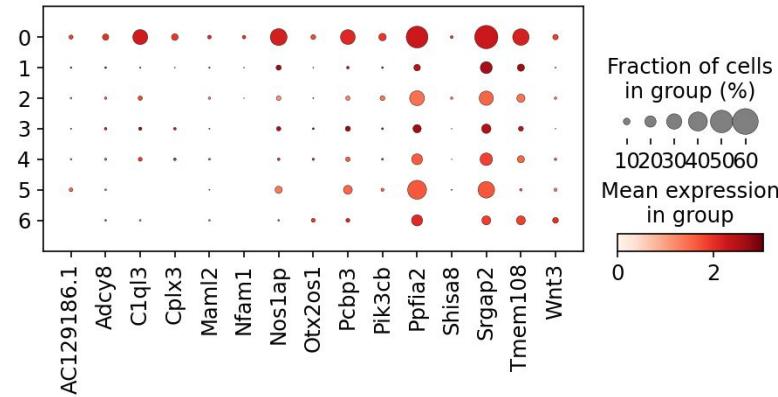


p0_p4_6
p0
Cluster 5



← 73.8% Correlation →

p0_p4_6
p4_6
Cluster 0

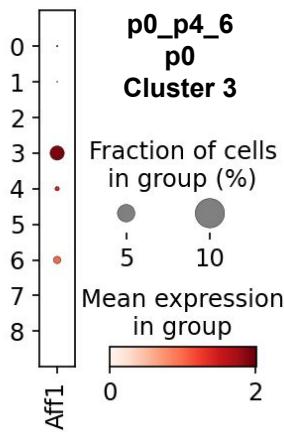
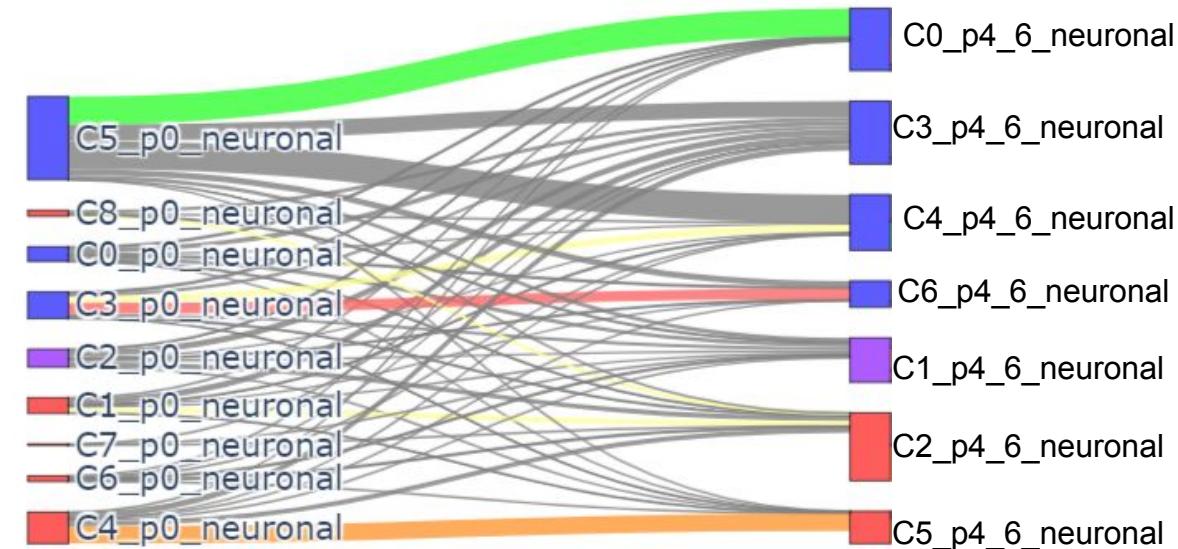


Fraction of cells
in group (%)
0 10 20 30 40 50 60

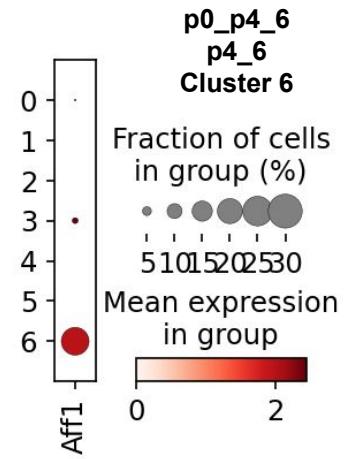
Mean expression
in group
0 2

C3_P0 vs C6_P4/6

Correlated Unique Genes

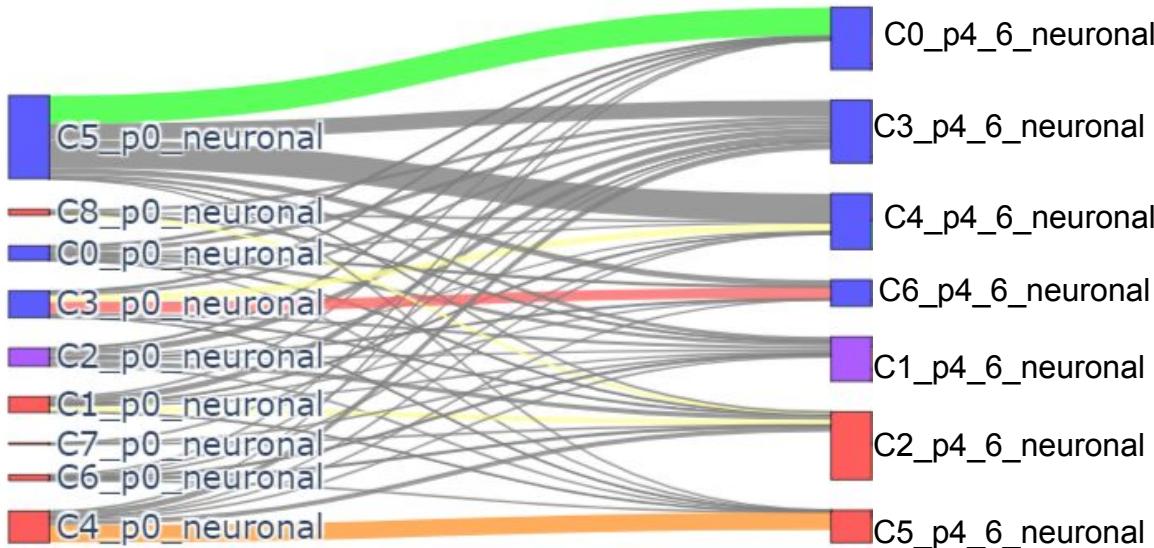


← 28.2% Correlation →



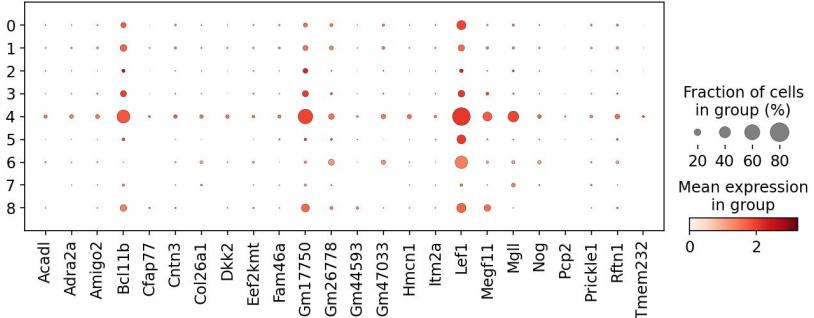
C4_P0 vs C5_P4/6 Correlated Unique Genes

Cntn3 - mediates layer specific targeting of retinal neurons
(Yamagata and Sanes, 2012)

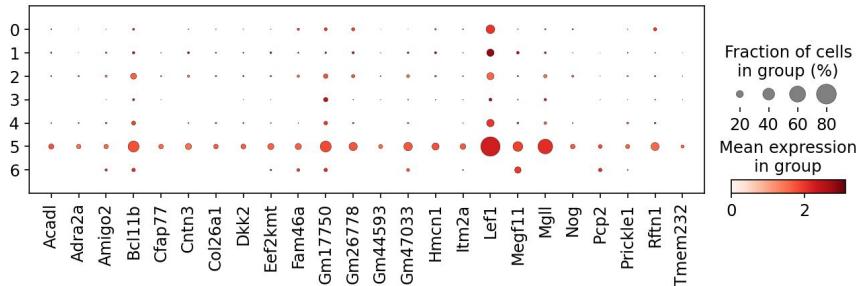


p0_p4_6
p0
Cluster 4

← 45.5% Correlation →



p0_p4_6
p4_6
Cluster 5



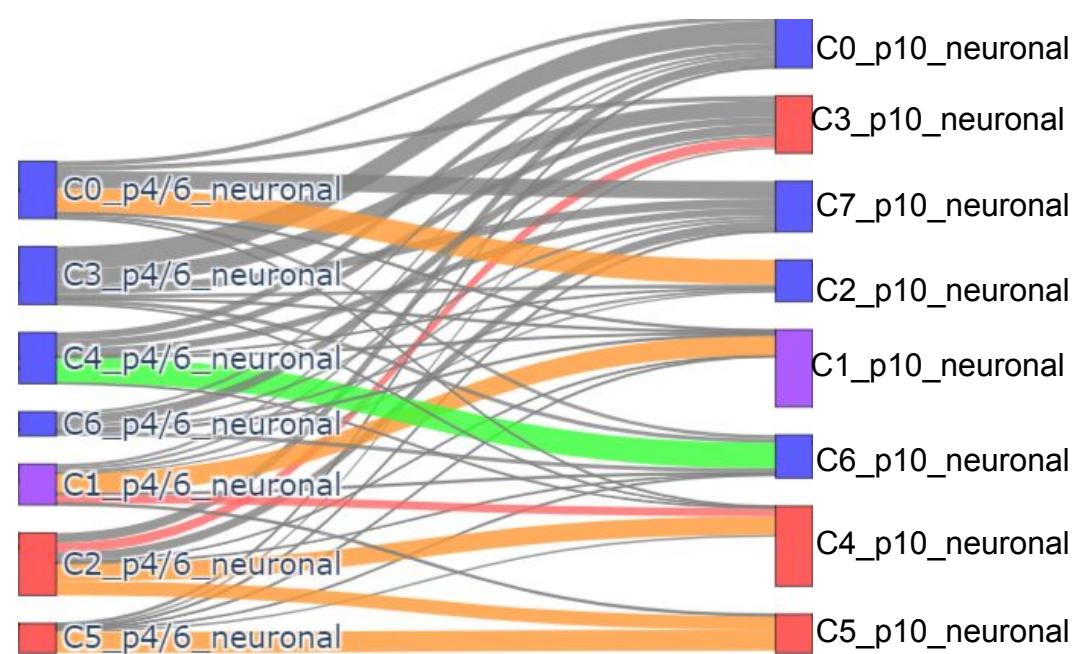
P4/6 vs P10

Correlated Unique Genes

C0_P4/6 vs C2_P10 Correlated Unique Genes

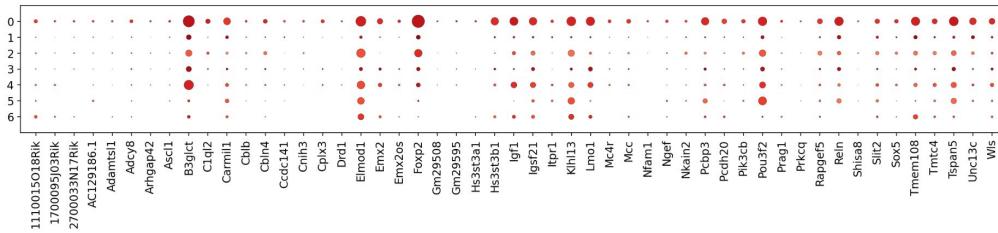
Pcdh20 - key cell adhesion molecule
Foxp2 - Corticofugal Neurons

(FOXP2 exhibits projection neuron class specific expression, but is not required for multiple aspects of cortical histogenesis, 2019)

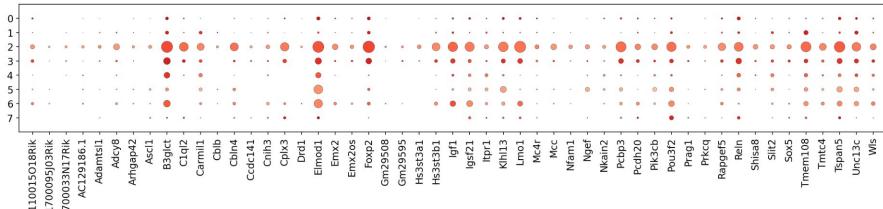


p4_6_p10
p4_6
Cluster 0

← 69.3% Correlation →

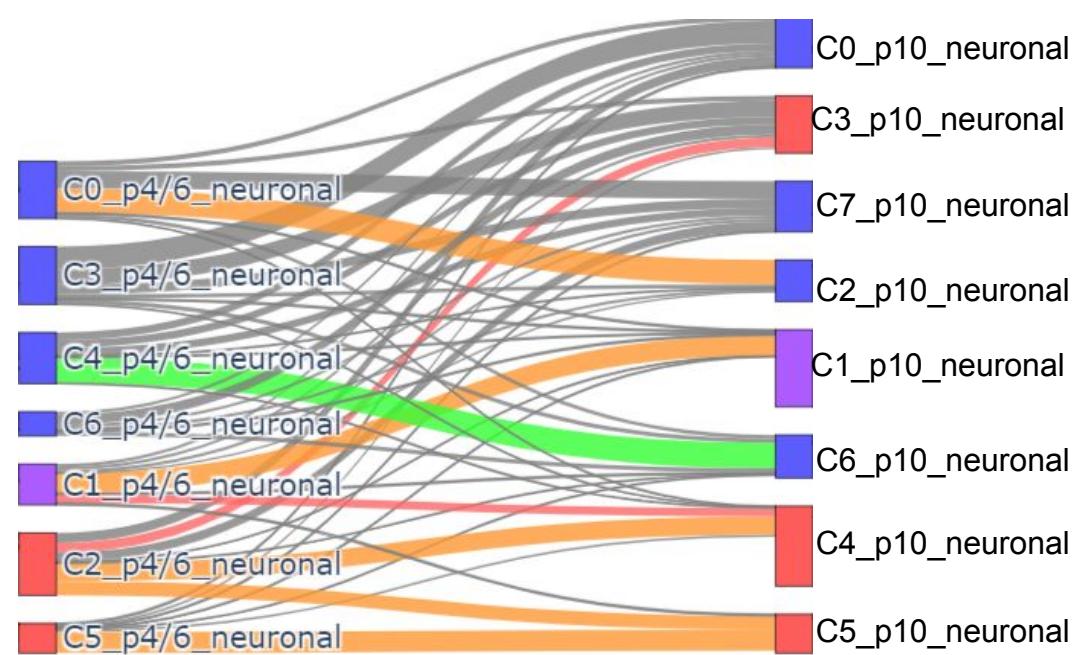


p4_6_p10
p10
Cluster 2



C4_P4/6 vs C6_P10

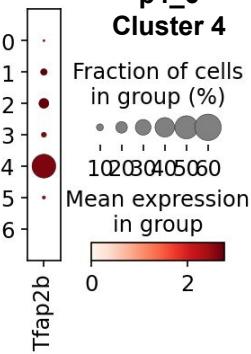
Correlated Unique Genes



p4_6_p10

p4_6

Cluster 4

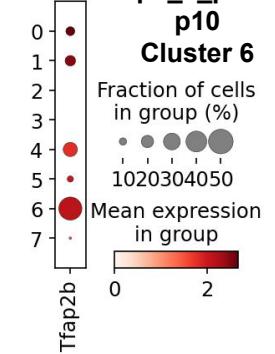


← 72.7% Correlation →

p4_6_p10

p10

Cluster 6



(Distinct Activities of Tfap2A and Tfap2B in the Specification of GABAergic Interneurons in the Developing Cerebellum, 2017)

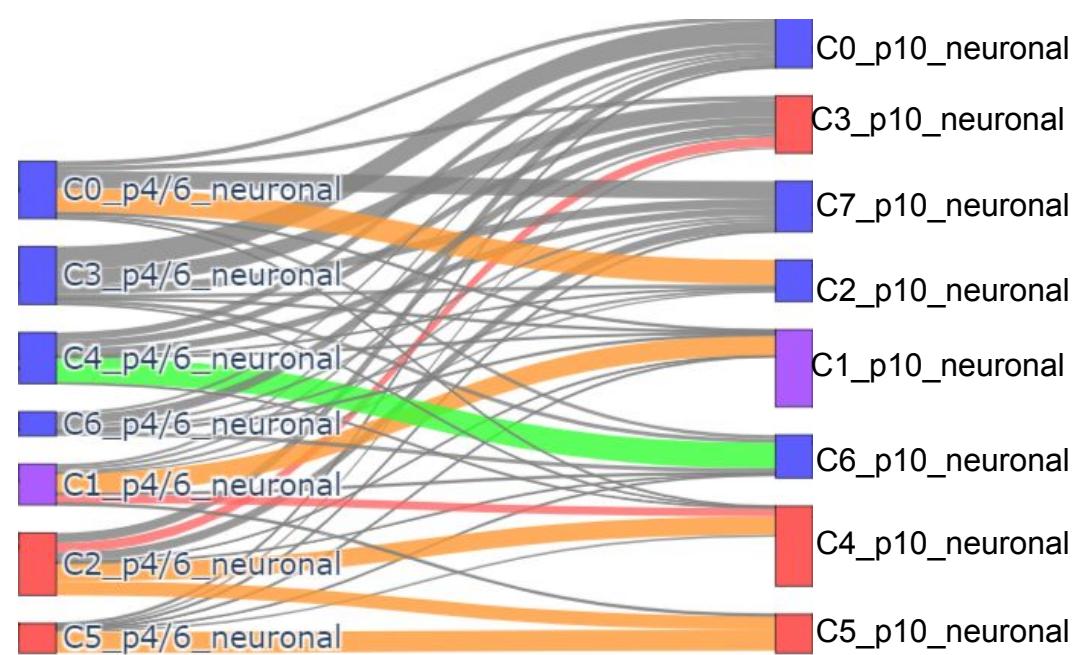
Tfap2b



- **Primarily selectively expressed in the interneurons**
- **Found heavily in the Intermediate/Deep layers of the SC**
- **Possibly links C4_P4/6 and C6_P10 to the Intermediate/Deep layers of the SC as well as linking its neuronal subtype to interneurons**

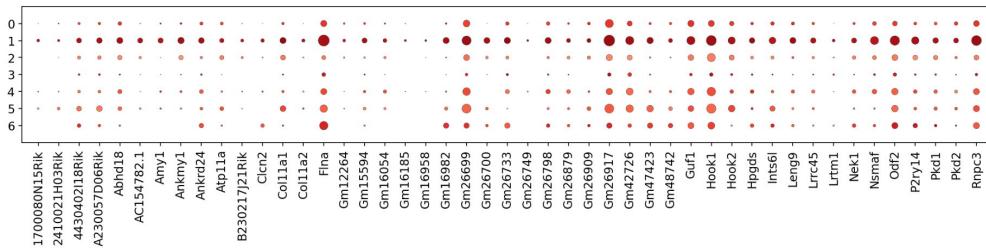
C1_P4/6 vs C1_P10 Correlated Unique Genes

Gm genes?

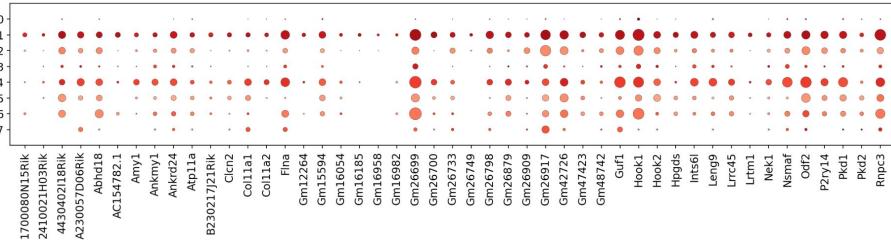


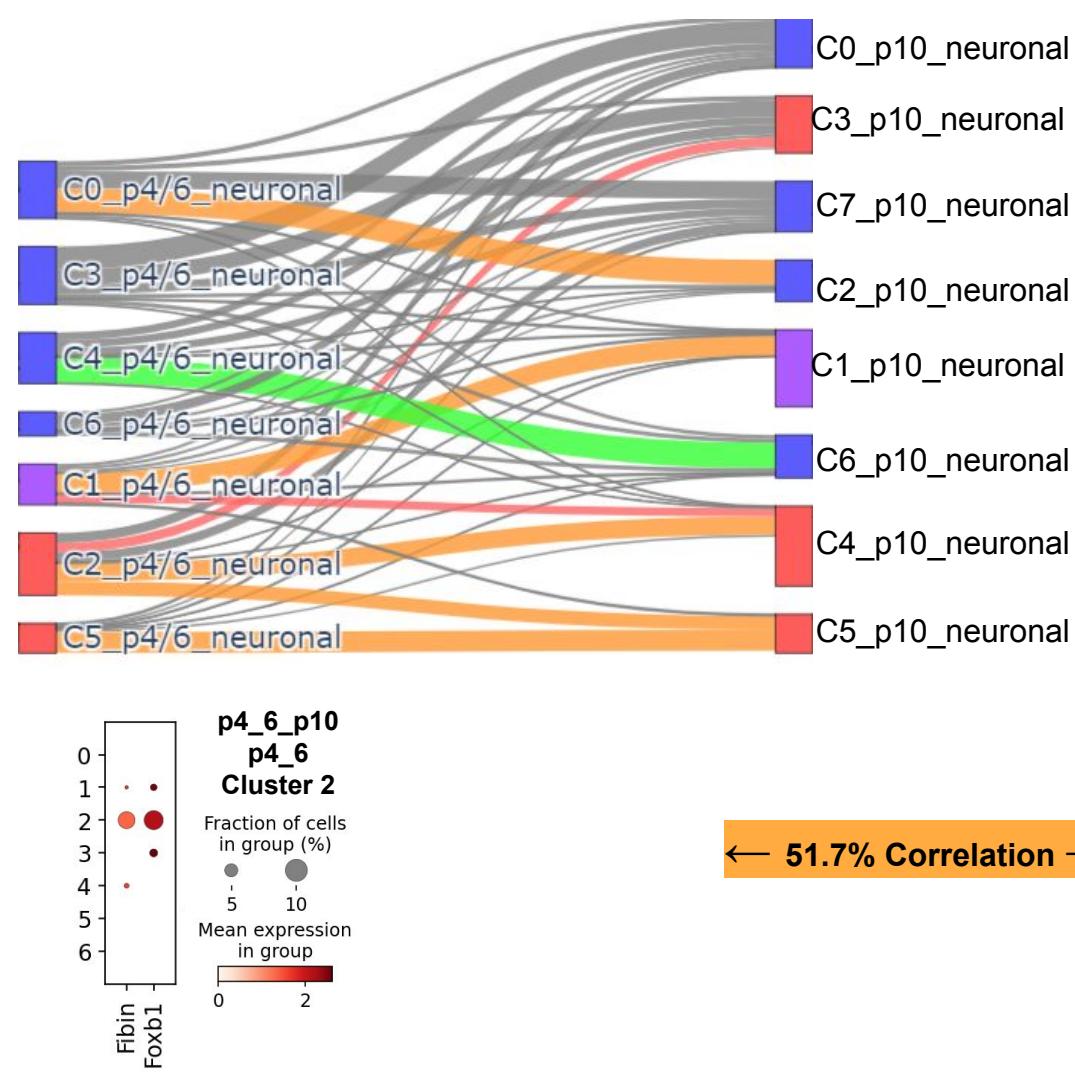
p4_6_p10
p4_6
Cluster 1

← 57.2% Correlation →



p4_6_p10
p10
Cluster 1





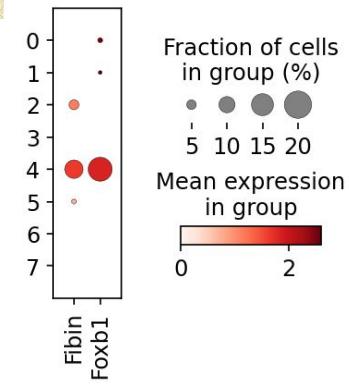
C2_P4/6 vs C4_P10

Correlated Unique Genes

Foxb1 - Found moderately within the Intermediate/Deep SC

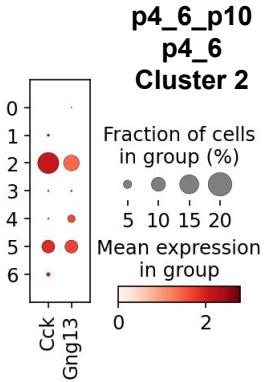
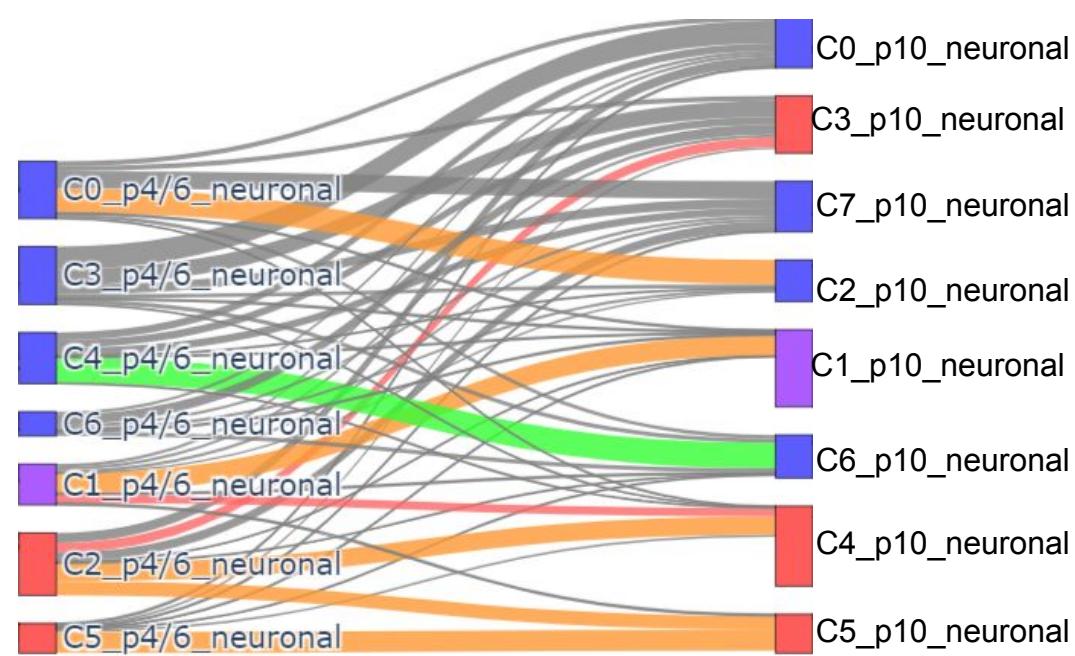


p4_6_p10
p10
Cluster 4

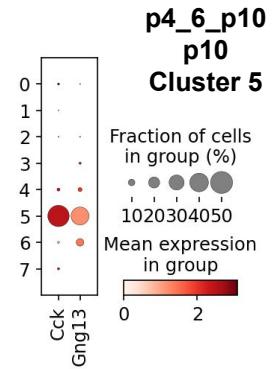


C2_P4/6 vs C5_P10

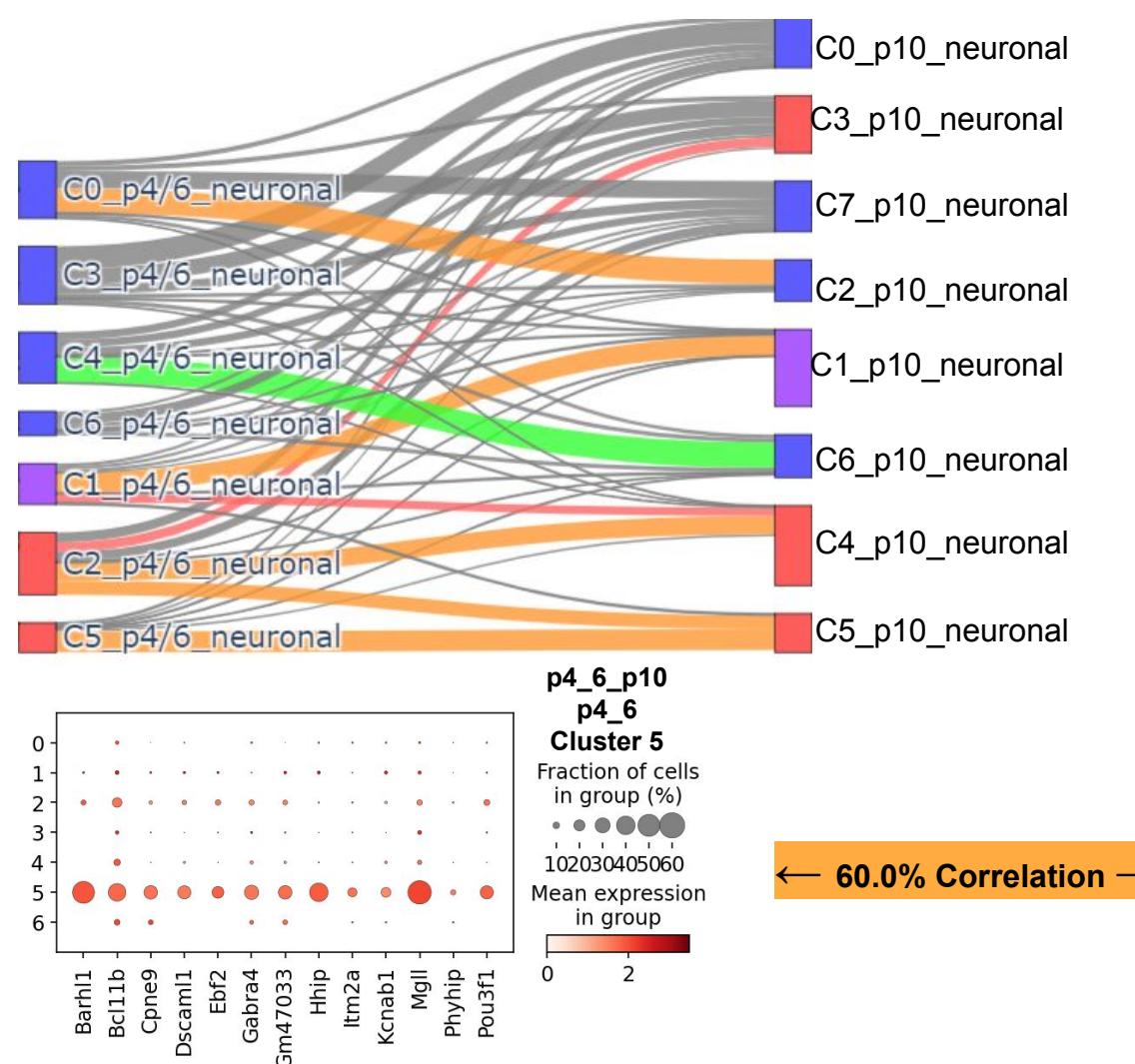
Correlated Unique Genes



← 40.0% Correlation →

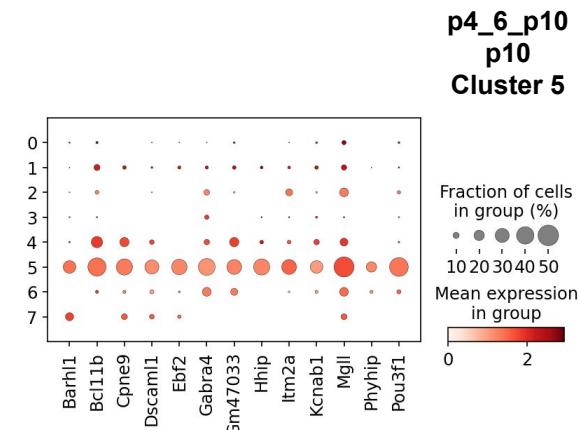


Cck - Cholecystokinin key Neuropeptide/Ca²⁺ binding protein
Linked to Cannabinoid Receptor 1



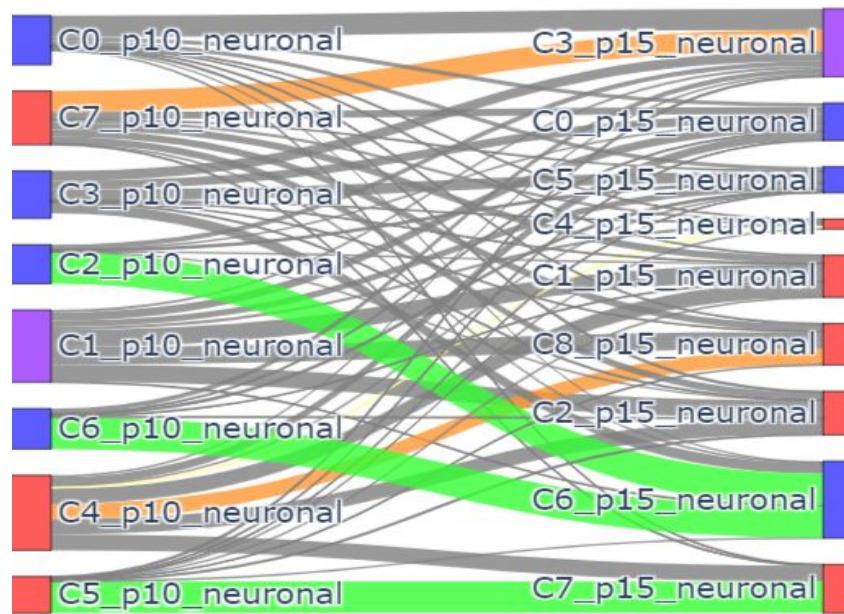
C5_P4/6 vs C5_P10 Correlated Unique Genes

BarH1 and Hhip- both are highly expressed within the superficial SC



P10 vs P15

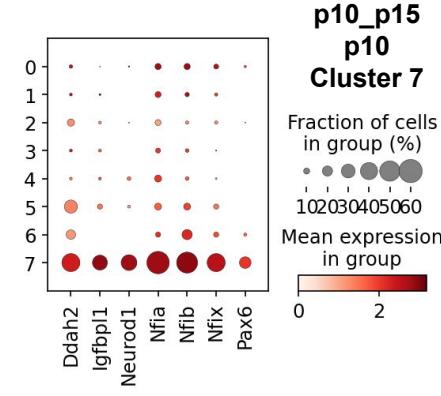
Correlated Unique Genes



C7_P10 vs C3_P15 Correlated Unique Genes

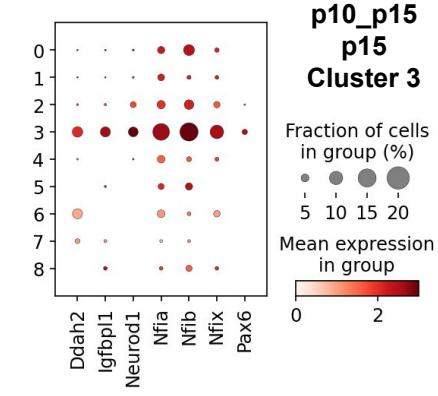
Pax6- key transcription factor
within the SC

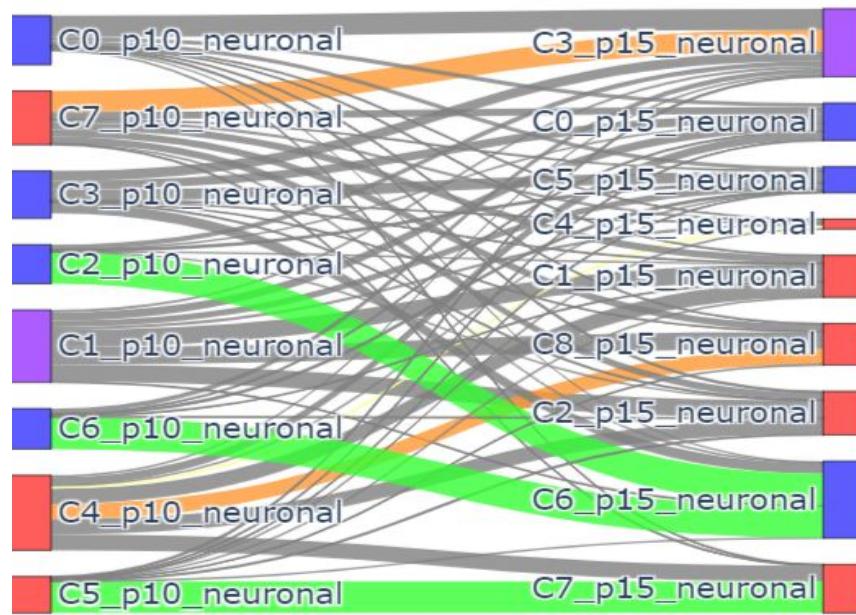
(*The Superior Colliculus: Cell Types, Connectivity, and Behavior*)



← 64.3% Correlation →

positive regulation of glutamatergic neuron differentiation GO:0120008
Pax6





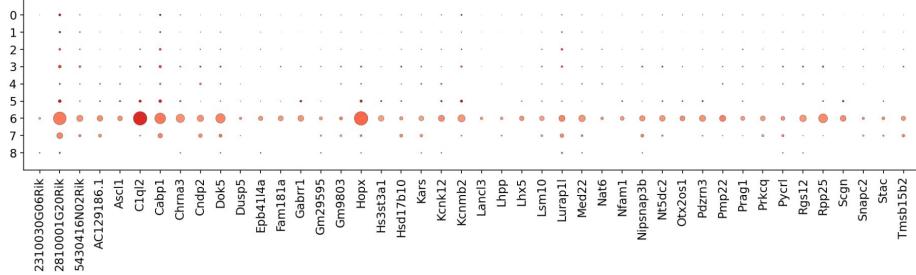
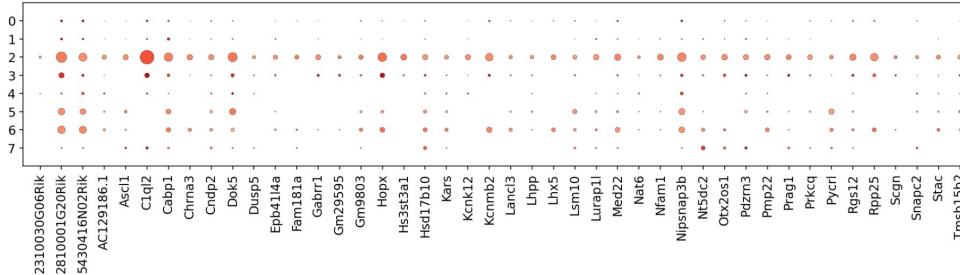
C2_P10 vs C6_P15 Correlated Unique Genes

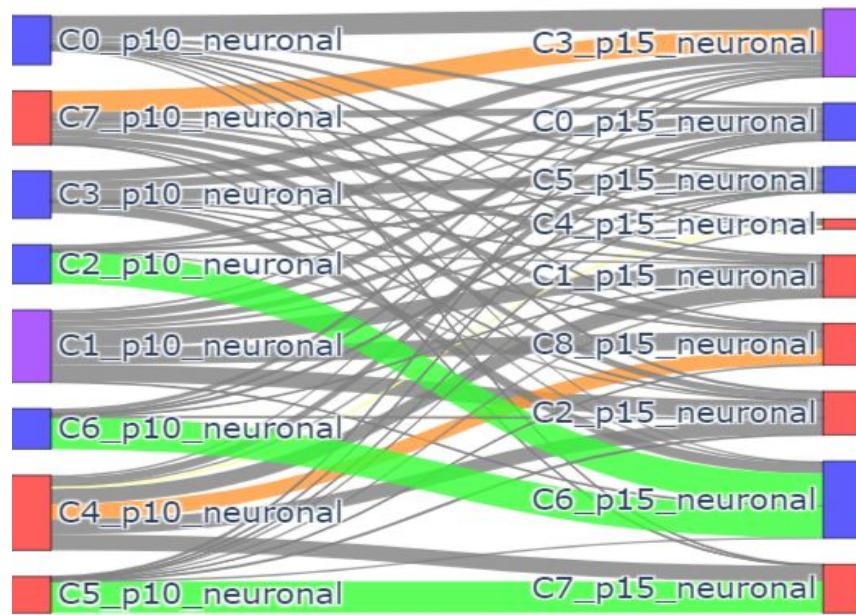
Postsynaptic nicotinic acetylcholine
receptors R-MMU-622327
Chrna3

p10_p15
p10
Cluster 2

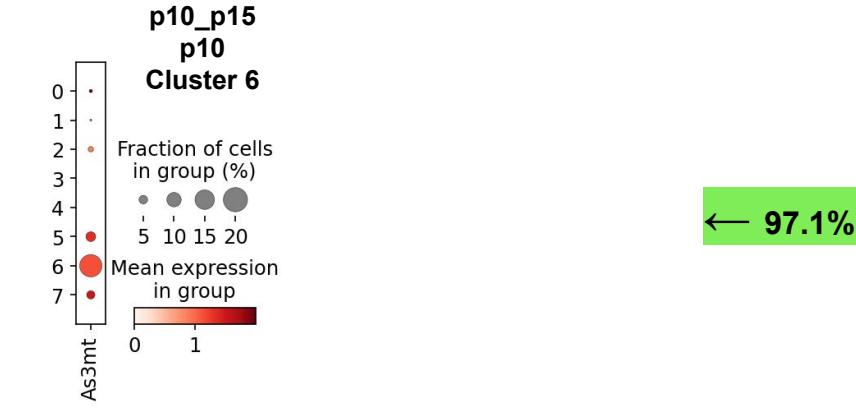
← 97.1% Correlation →

p10_p15
p15
Cluster 6

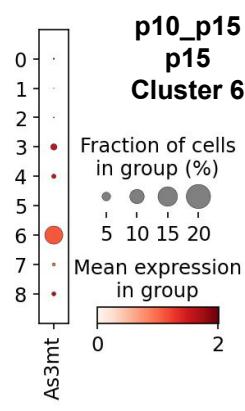


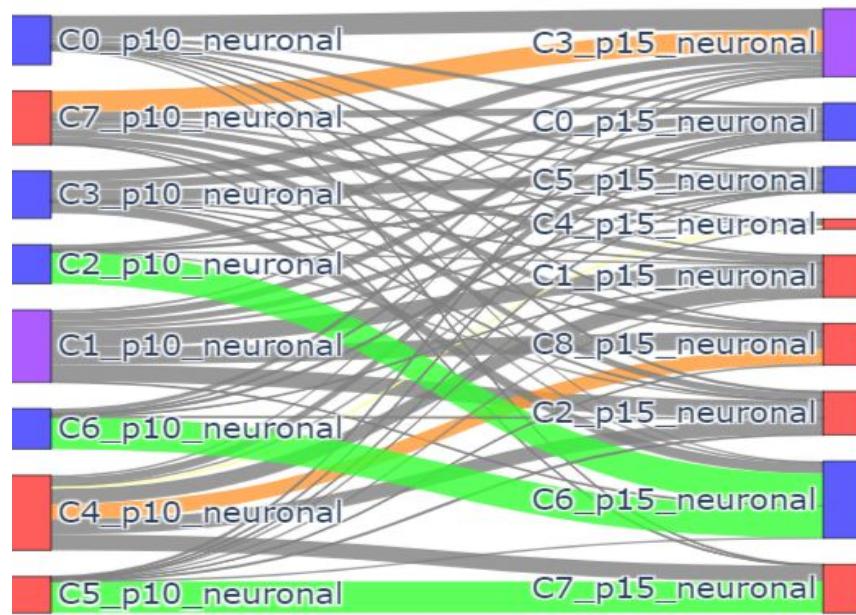


C6_P10 vs C6_P15 Correlated Unique Genes



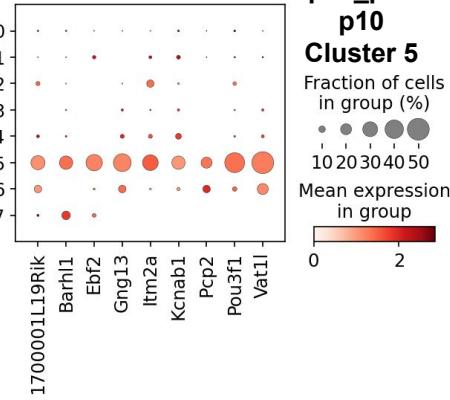
← 97.1% Correlation →





**p10_p15
p10
Cluster 5**

Fraction of cells
in group (%)
• • • 10 20 30 40 50
Mean expression
in group

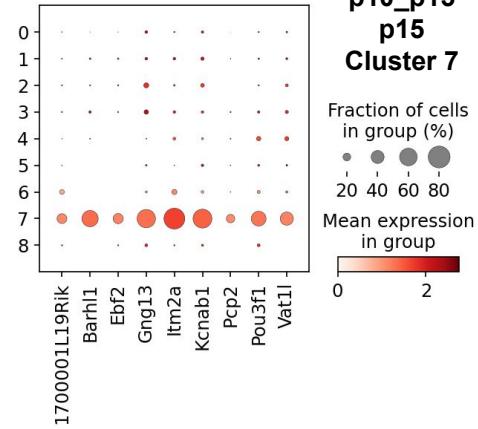


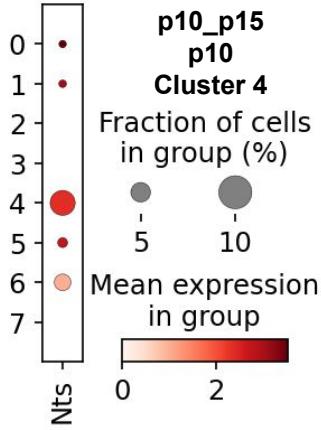
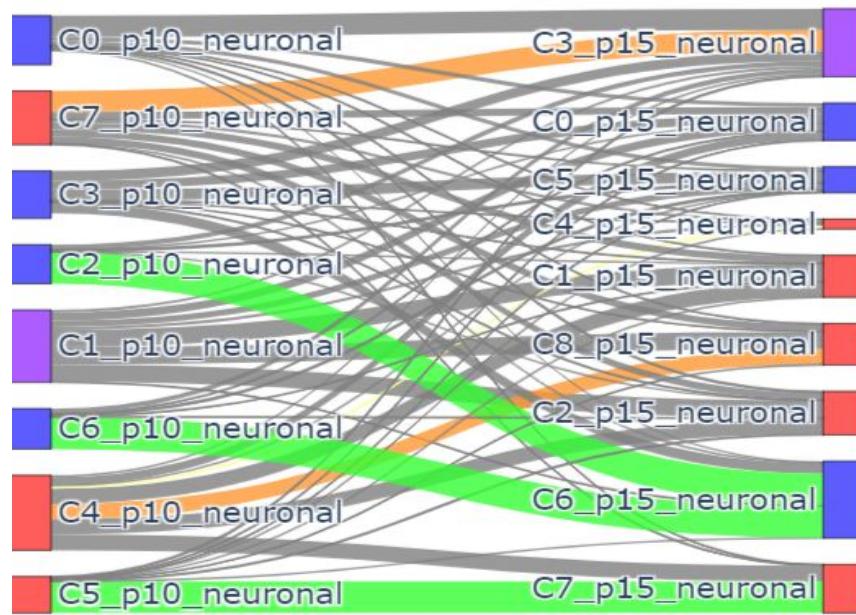
← 96.0% Correlation →

C5_P10 vs C7_P15 Correlated Unique Genes

**p10_p15
p15
Cluster 7**

Fraction of cells
in group (%)
• • • 20 40 60 80
Mean expression
in group

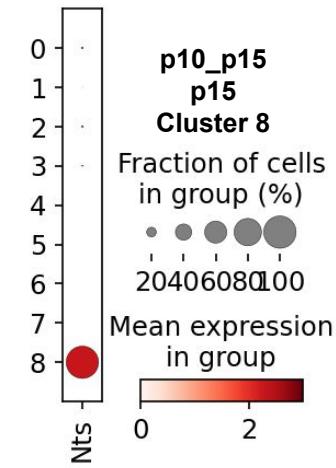




← 46.6% Correlation →

C4_P10 vs C8_P15 Correlated Unique Genes

Nts - Neurotensin 1 key receptor in the SC
(*The Superior Colliculus: Cell Types, Connectivity, and Behavior*)



Does High Cluster Correlation Indicate Unique Gene Enrichment?

P0_P4/6	Number of Genes	Correlation Percentage
Cluster_1_Cluster_2	10	12.9
Cluster_3_Cluster_6	1	28.2
Cluster_4_Cluster_5	24	45.5
Cluster_5_Cluster_0	15	73.8
Cluster_7_Cluster_0	1	0
Cluster_7_Cluster_6	2	0
Cluster_8_Cluster_2	3	2.86
Cluster_3_Cluster_4	1	17.1

P4/6_P10	Number of Genes	Correlation Percentage
Cluster_0_Cluster_2	51	69.3
Cluster_1_Cluster_1	47	57.2
Cluster_2_Cluster_5	2	40
Cluster_5_Cluster_5	13	80
Cluster_2_Cluster_4	2	61.7
Cluster_2_Cluster_7	1	28
Cluster_4_Cluster_6	2	72.7
Cluster_1_Cluster_4	8	24.7

P10_p15	Number of Genes	Correlation Percentage
Cluster_2_Cluster_6	44	97.1
Cluster_5_Cluster_7	9	88
Cluster_7_Cluster_3	7	64.3
Cluster_6_Cluster_6	1	97.1
Cluster_4_Cluster_8	1	48.8
Cluster_4_Cluster_4	7	8.78

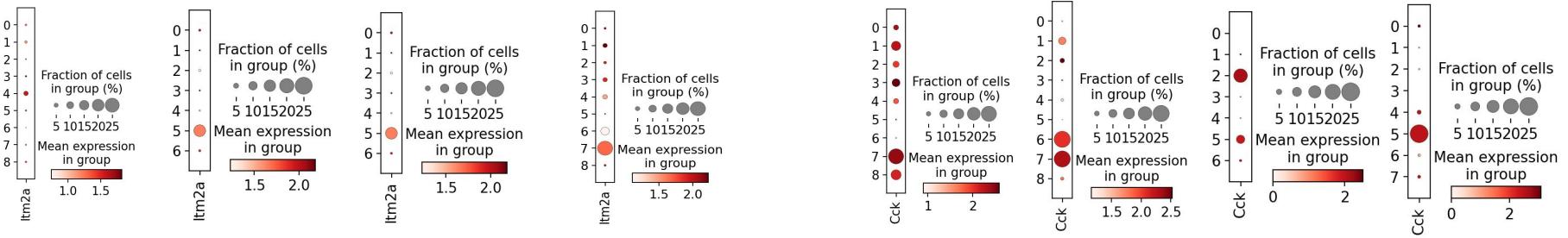
Average Gene Number: Correlation >70	15.5
Average Gene Number: Correlation 40-70%	18.1
Average Gene Number: Correlation 20-40%	3.3
Average Gene Number: Correlation 0-20%	2.8

C0_P0_C2_P4/6 → 69.3%
Correlation at Cutoff

Adjusted

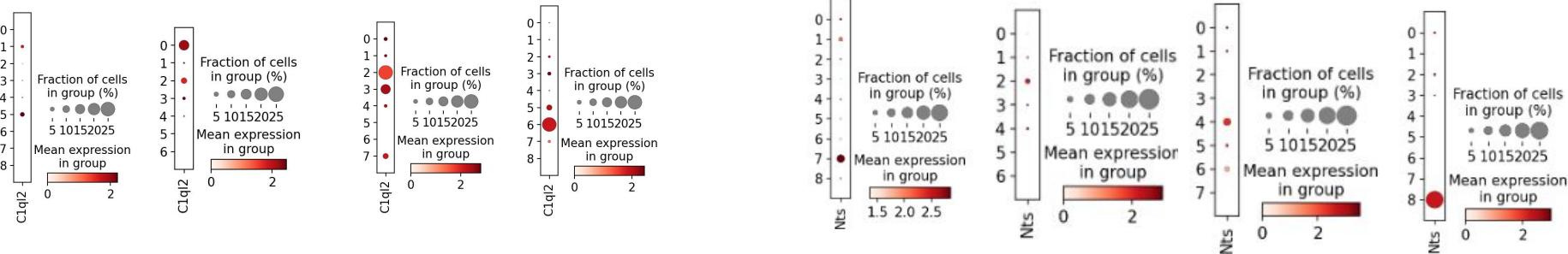
Average Gene Number: Correlation >70	19.2
Average Gene Number: Correlation 40-70%	16.6
Average Gene Number: Correlation 20-40%	3.3
Average Gene Number: Correlation 0-20%	2.8

Up/Down Regulation of Correlated Unique Genes



P0 → P15

P15



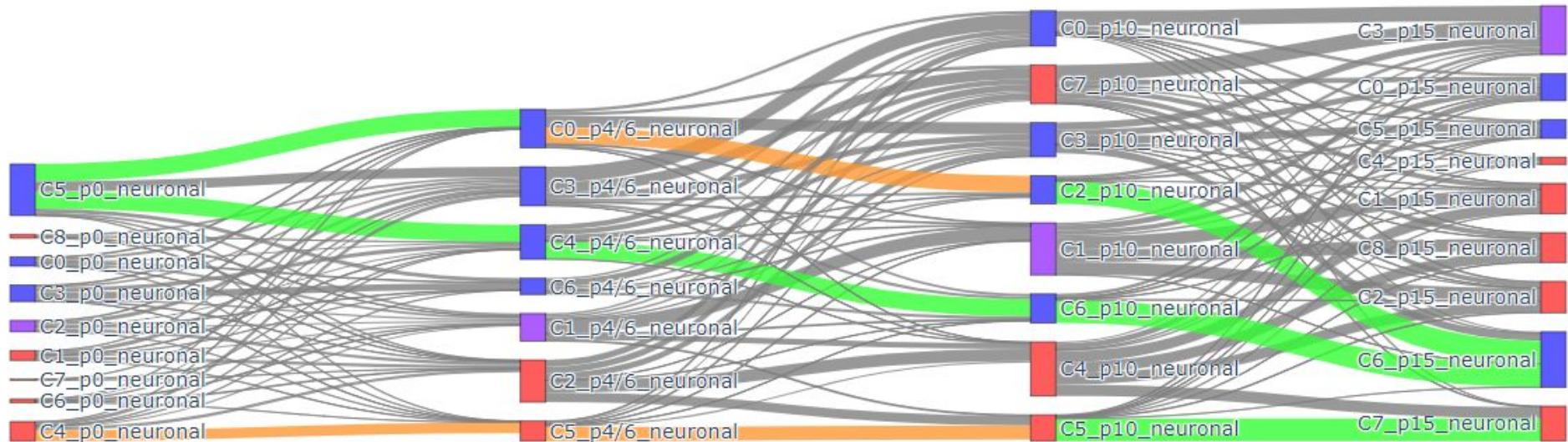
P0 → P15

P15

Conserved Clusters Through Time

P0 vs P4/6 vs P10 vs P15

All Conserved Clusters

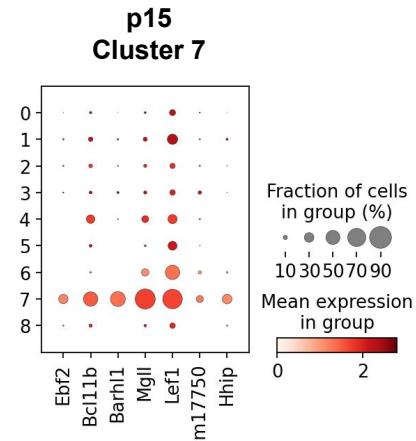
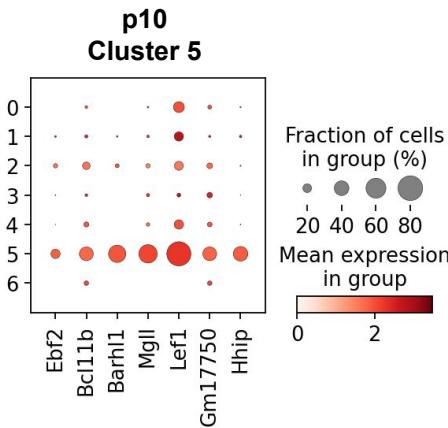
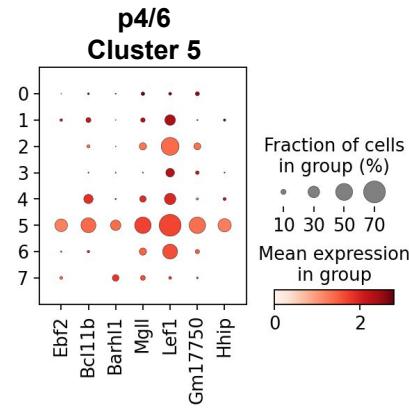
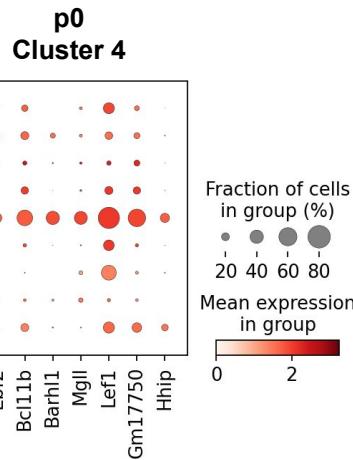
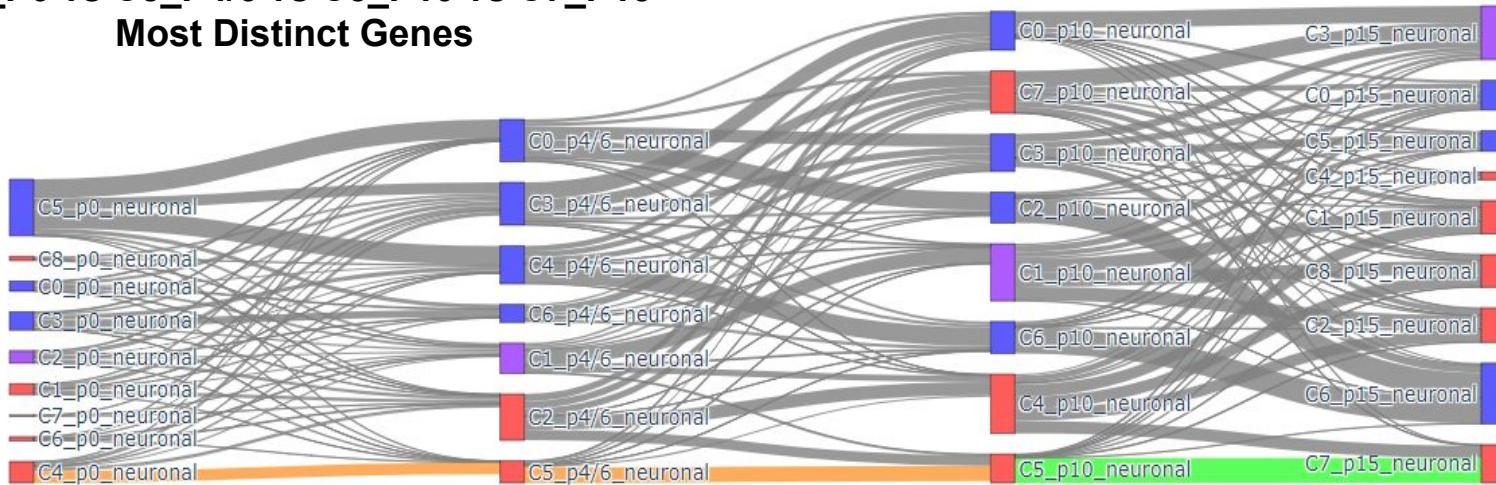


C4_P0 vs C5_P4/6 vs C5_P10 vs C7_P15

C5_P0 vs C0_P4/6 vs C2_P10 vs C6_P15

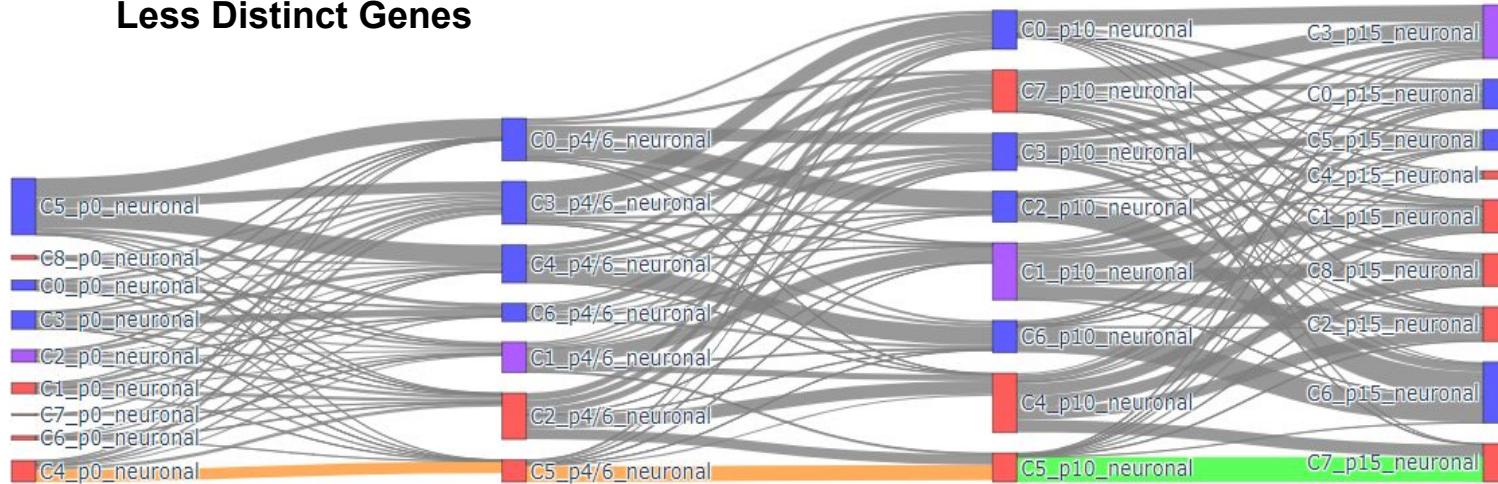
C5_P0 vs C4_P4/6 vs C6_P10 vs C6_P15

C4_P0 vs C5_P4/6 vs C5_P10 vs C7_P15 Most Distinct Genes

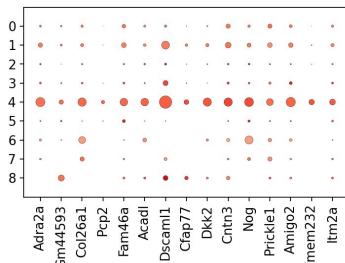


C4_P0 vs C5_P4/6 vs C5_P10 vs C7_P15

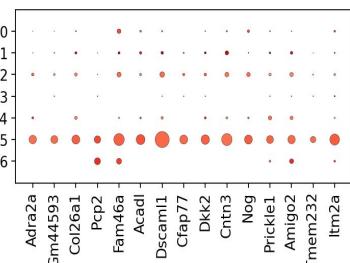
Less Distinct Genes



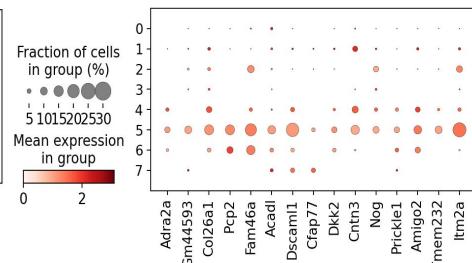
p0
Cluster 4



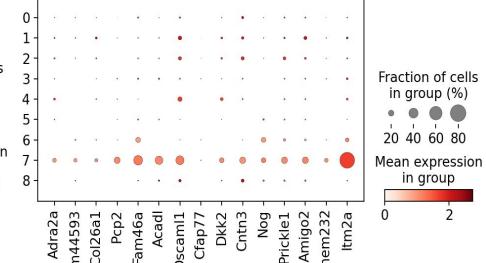
p4/6
Cluster 5



p10
Cluster 5



p15
Cluster 7



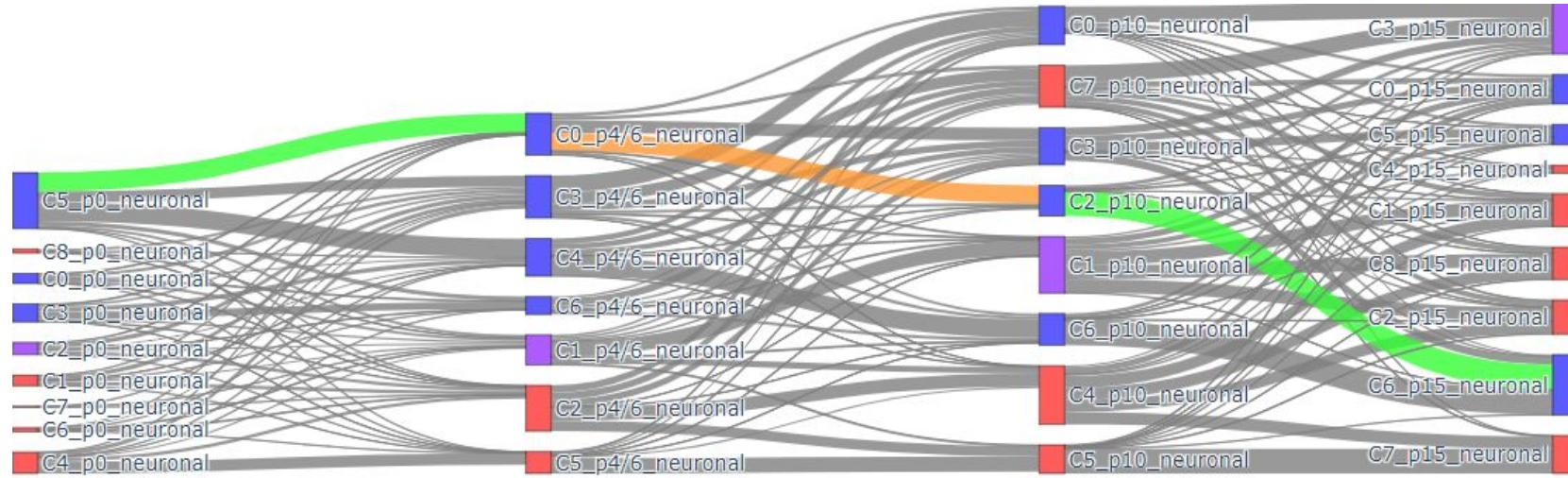
Barh1 and Hhip



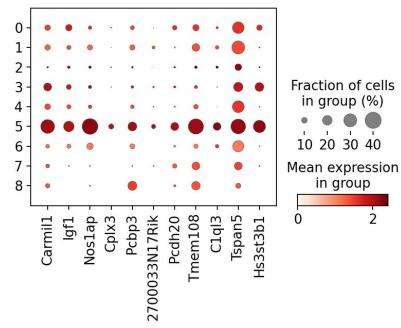
- Found heavily in the Superficial SC
- Possibly links C4_P0, C5_P4/6, C5_P10, C7_P15 to the Superficial layers of the SC as well possibly linking other correlated genes to the Superficial layer

C5_P0 vs C0_P4/6 vs C2_P10 vs C6_P15

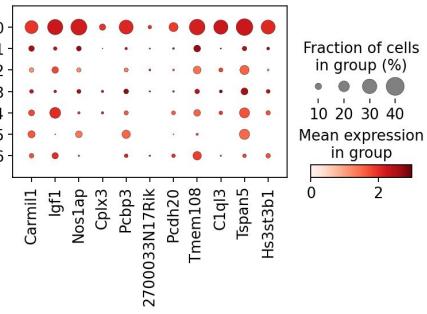
Distinct Genes



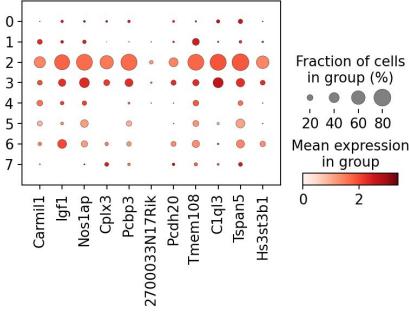
p0
Cluster 5



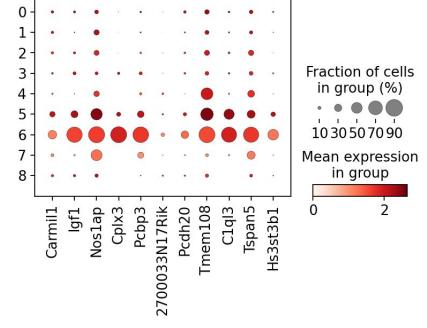
p4/6
Cluster 0



p10
Cluster 2

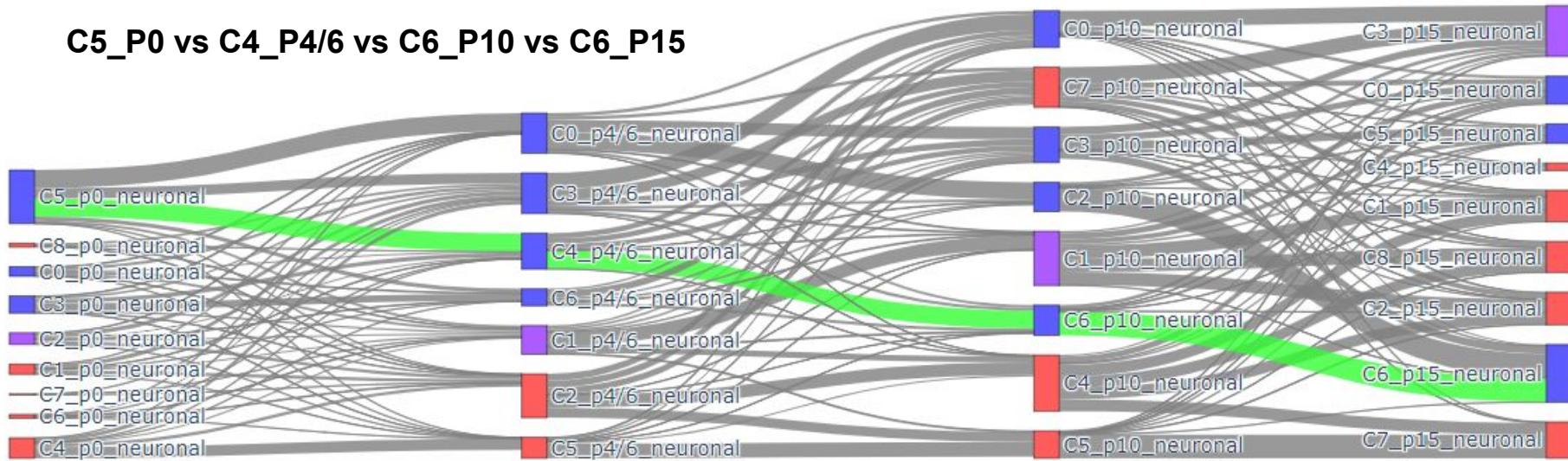


p15
Cluster 6



Exception:

C5_P0 vs C4_P4/6 vs C6_P10 vs C6_P15



Next Steps

- Visualize GO enrichment through the use of GOAtools/GOpplot to provide functional names for genes and clusters
- Identify and find ways to visualize which genes within correlated clusters are being up regulated and down regulated through time
- Read through the literature and link more unique/SD genes to different layers within the SC

End Goal:

- Create a complete mapping of clusters and their genes through time, the functional elements that change, and their corresponding links to regions within the layers of the SC.