

scRNA analysis on *Mus musculus* Spinal cord tissue

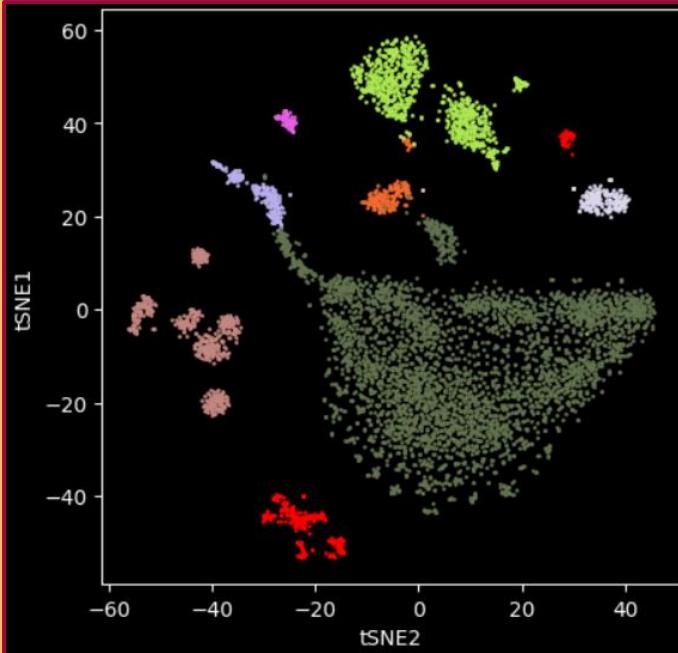
BCG 2023 – 2024
Genomics and Transcriptomics course

Flavia Leotta



THE DATASET

SRA	SRS	TOT CELLS
SRA667466	SRS3060017	6,372



LINK TO PANGLAODB

https://panglaodb.se/view_data.php?sra=SRA667466&srs=SRS3060017#google_vignette

LINK TO THE ARTICLE OF THE STUDY

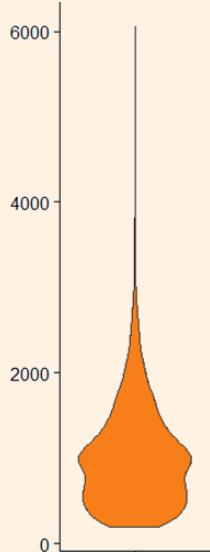
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10366611/>

DATA BY PANGALOBD

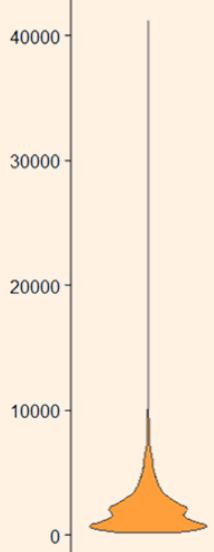
CELL TYPES	DESIGN
<ul style="list-style-type: none">Astrocytes (n = 187)Endothelial cells (n = 818)Ependymal cells (n = 74)Microglia (n = 187)Oligodendrocyte progenitor cells (n = 220)Oligodendrocytes (n = 3663)Trigeminal neurons (n = 553)Unknown (n = 348)	<ul style="list-style-type: none">26,764 expressed genesMedian number of expressed genes per cell = 113825 clusters7 cell types

QUALITY CONTROL

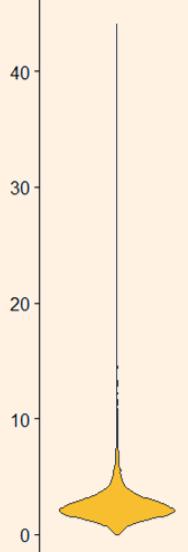
nFeature
RNA



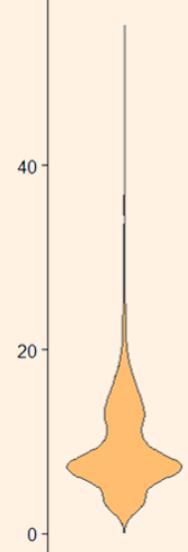
nCount
RNA



Percent
mtRNA



Percent
rRNA



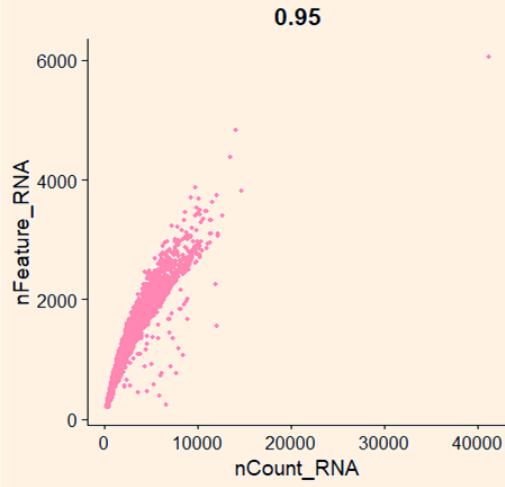
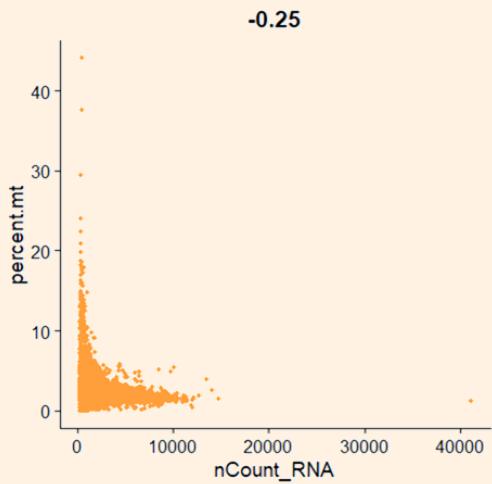
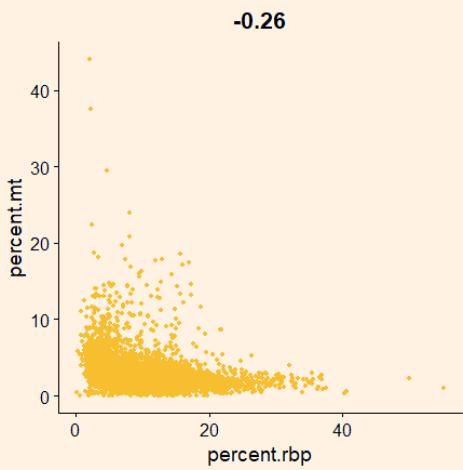
9089 samples

Filter out:

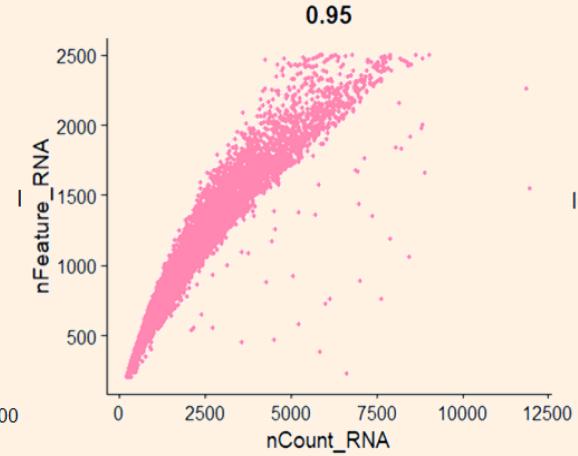
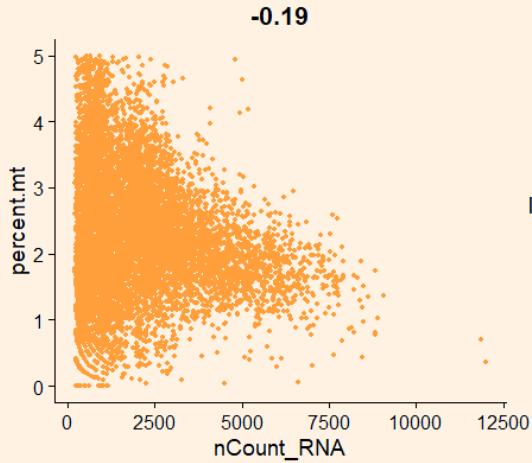
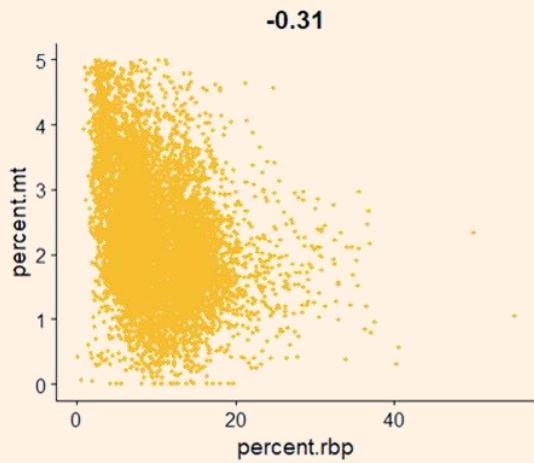
- Counts \leq 200
- Counts $>$ 2500
- %mt $>$ 5

8342 samples

CORRELATION pre-filtering

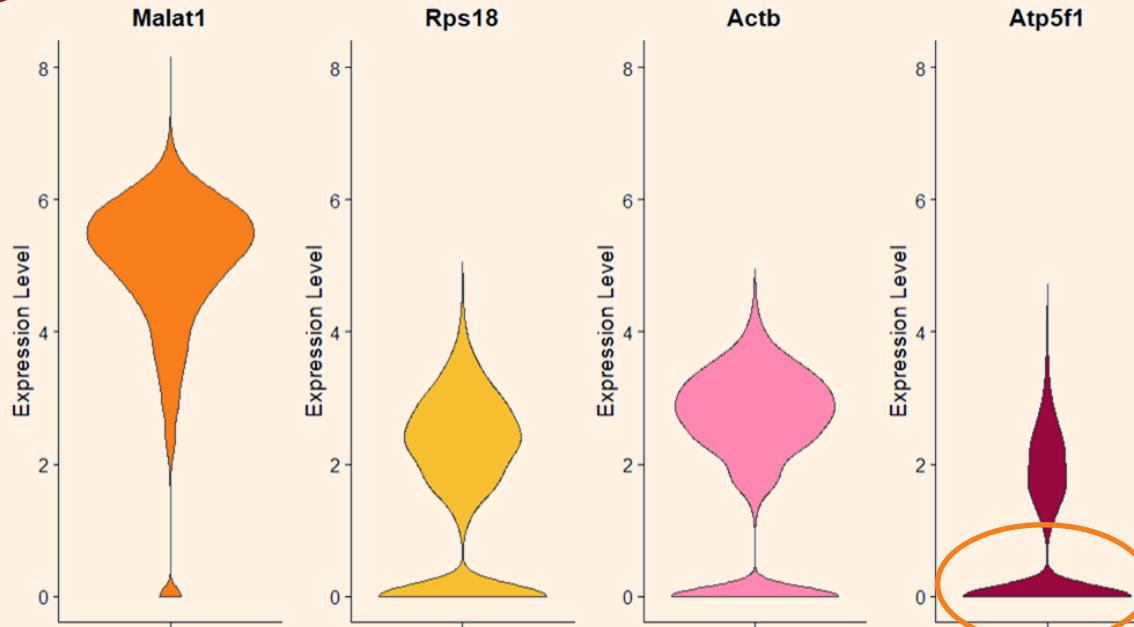


CORRELATION after filtering



MOST EXPRESSED VS HOUSEKEEPING*

AFTER NORMALIZATION

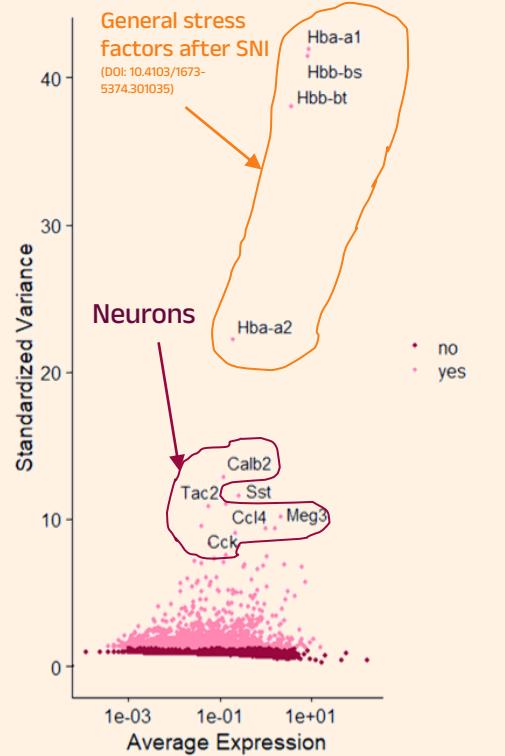


*Housekeeping genes selected from Yulia Panina *et al.* (2028)
doi.org/10.1038/s41598-018-26707-8

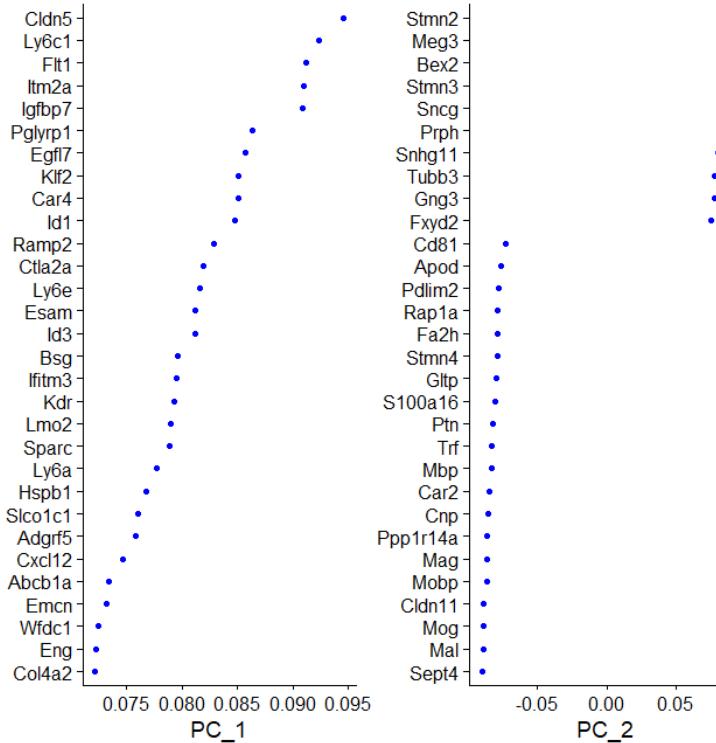
DROPOUTS

UP-REGULATED GENES

First 2k most upregulated genes

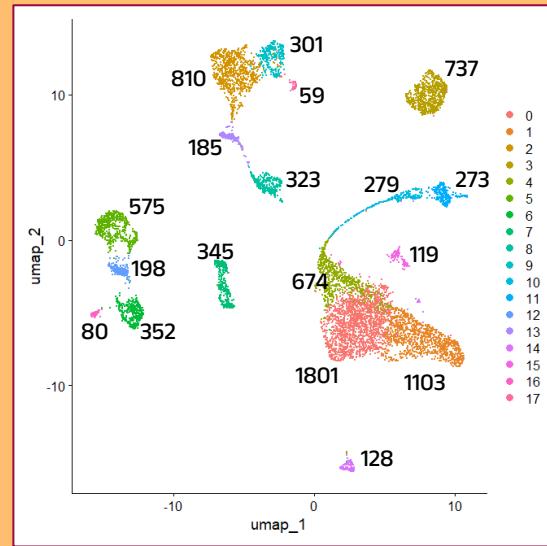
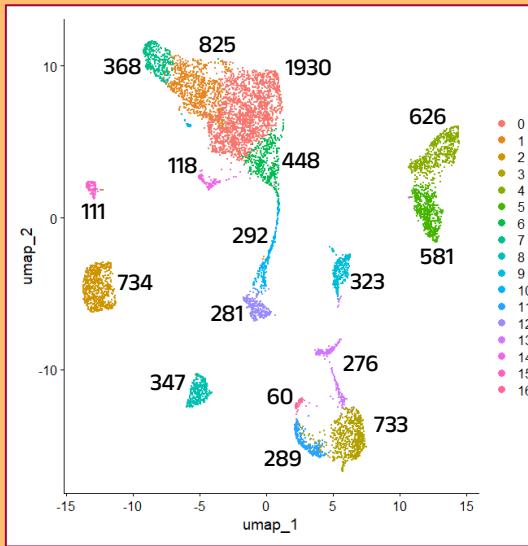
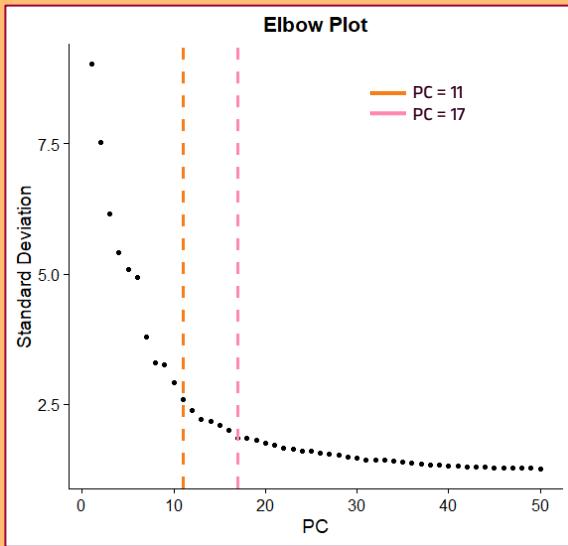


Most variable genes for the first 2 components



- **Cldn5:** marker for endothelial cells (Jonas C. Schupp et al., 2021)
- **Ly6c1:** possible sign of infection in endothelial tissues (Yuan-hui Li et al., 2022)
- **Flt1:** endothelial cells (panglaodb)
- **Stmn2:** Neurons (panglaodb)
- **Meg3:** Neurons (panglaodb)
- **Bex2:** Neurons (panglaodb)

PCA AND CLUSTERING

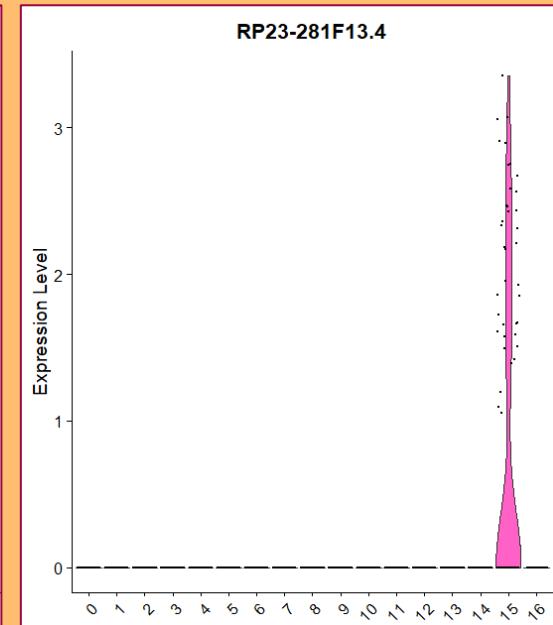
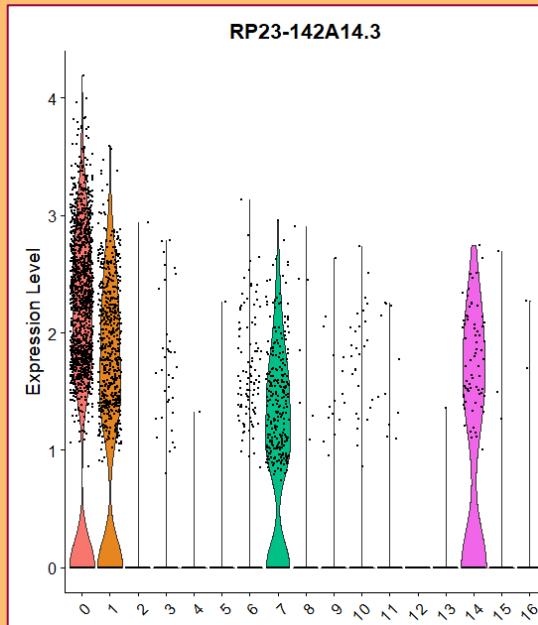
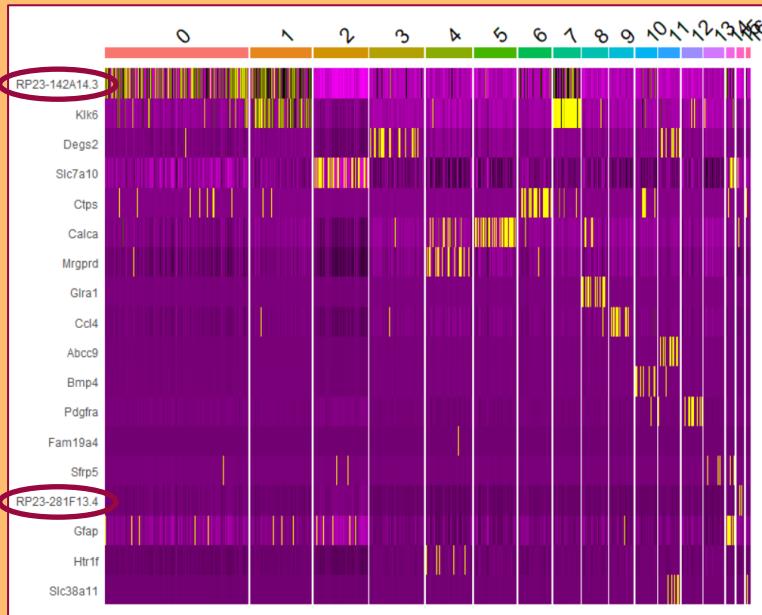


Number of PCs that explains 75% of the variance = 11.
The analysis was repeated also with number of PCs = 17.

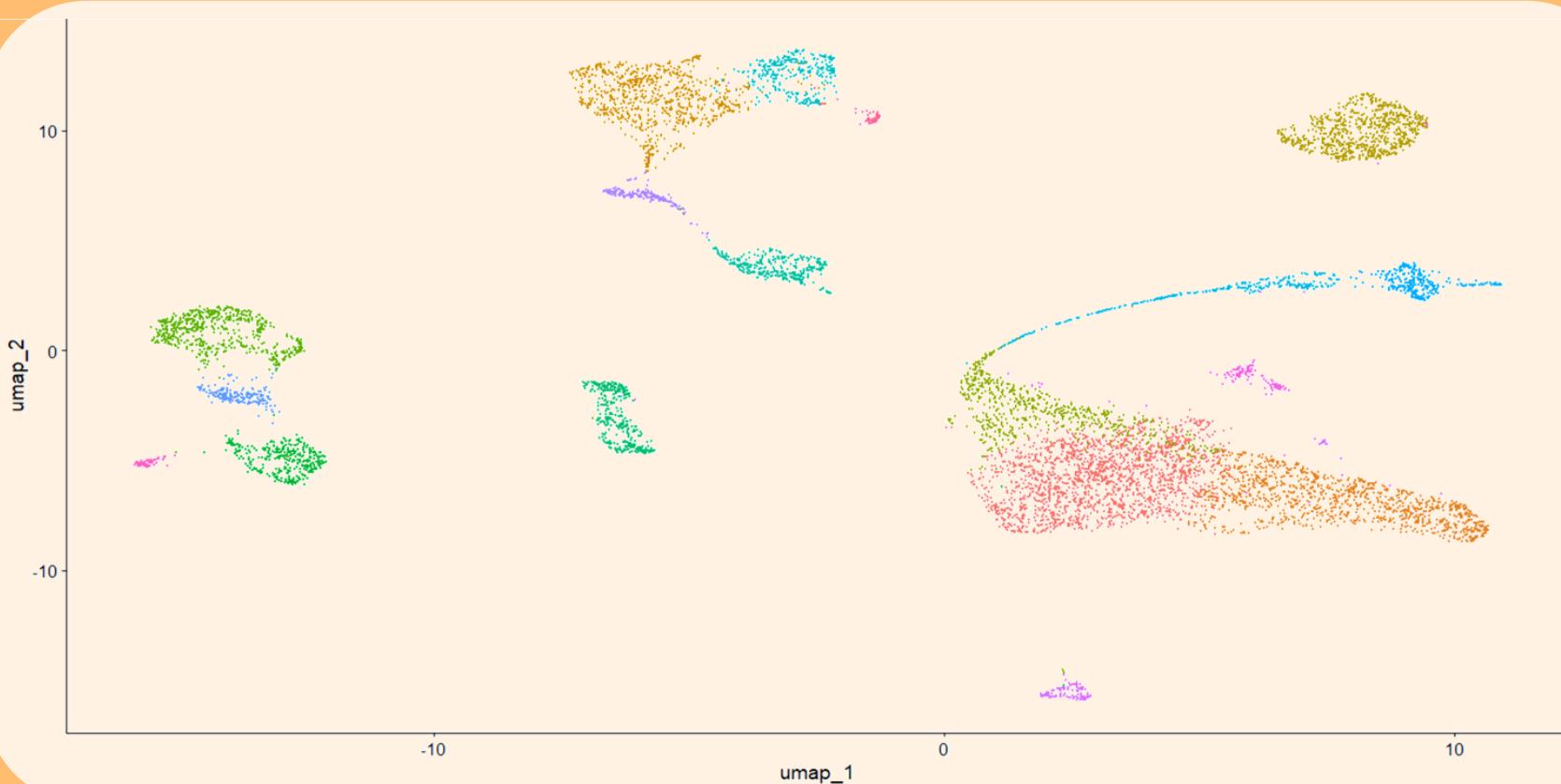
Using PC = 11 resulted in the formation of 17 clusters.
Now showing UMAP representation.

Using PC = 17 resulted in the formation of 18 clusters.
Now showing UMAP representation.

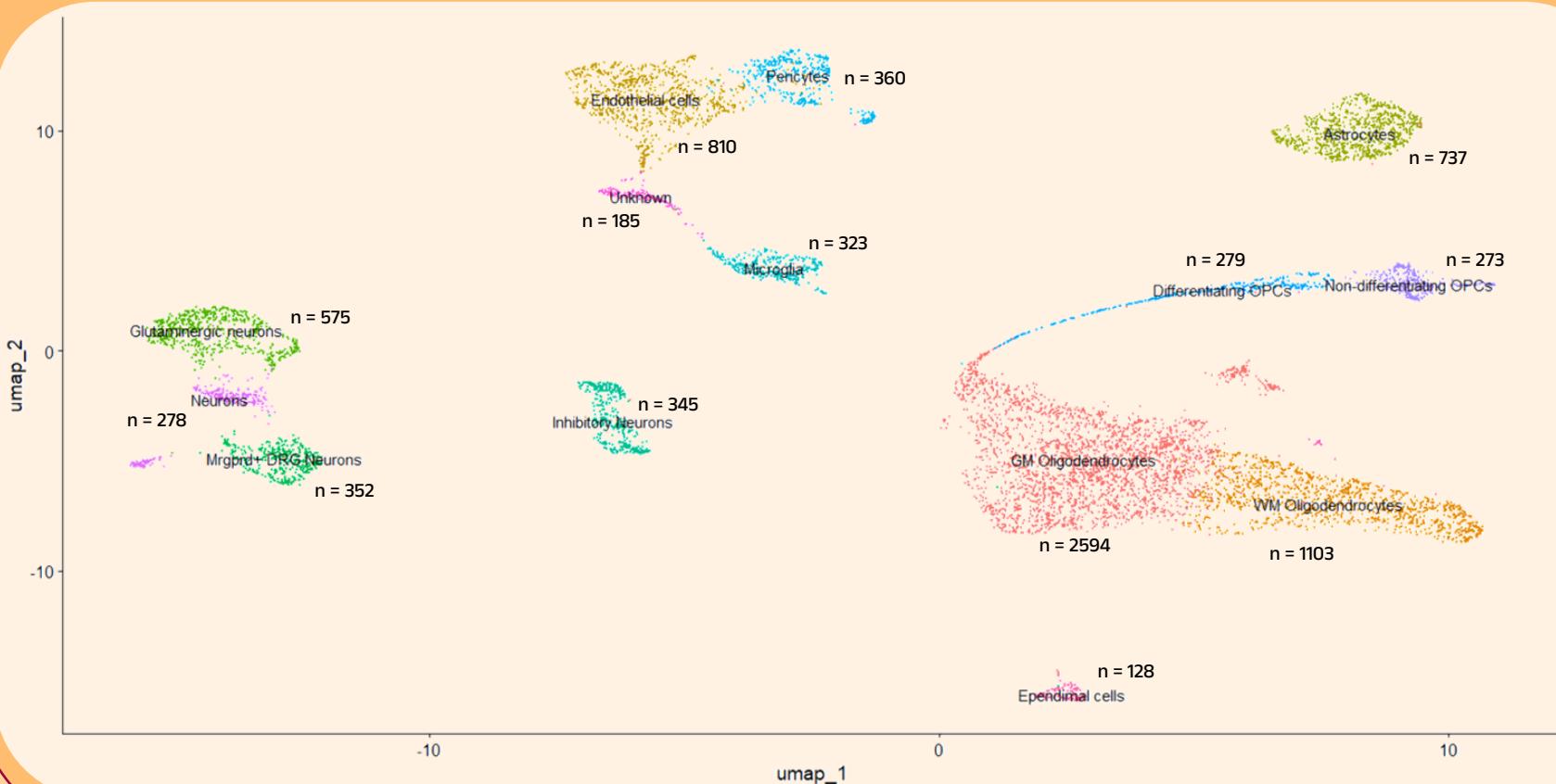
IS CLUSTERING DUE TO BIAS?



BEFORE



AFTER

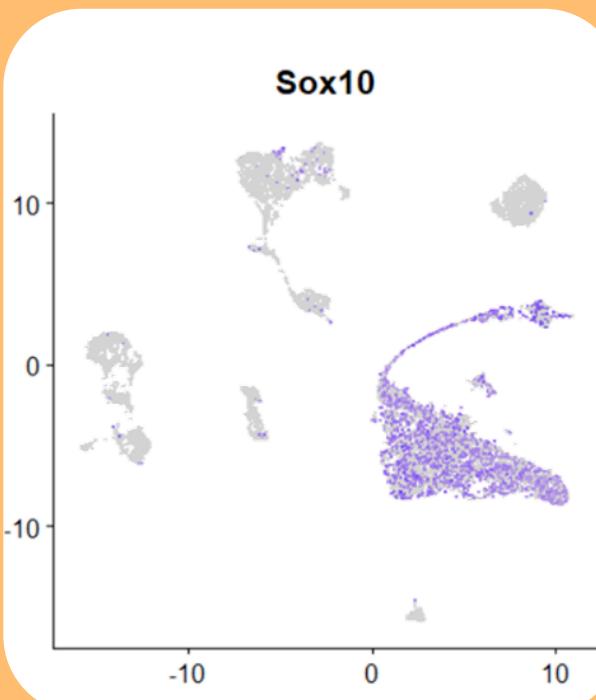


GENES AND CELL TYPES

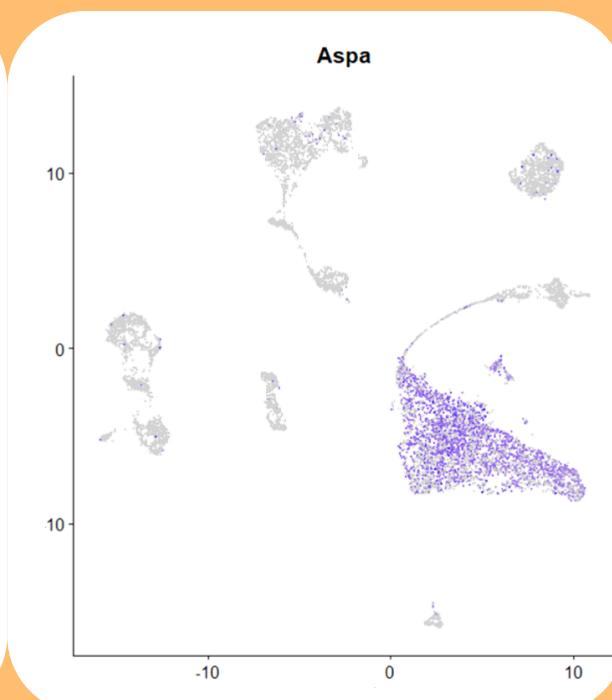
#	MARKER GENE(S)	CELL TYPE	SOURCE	#	MARKER GENE(S)	CELL TYPE	SOURCE
0	«Ptgds»	GM Oligodendrocytes	<ul style="list-style-type: none"> Elisa M. Floriddia <i>et al.</i> (2020) doi.org/10.1038/s41467-020-19453-x 	9	«Abcc9», «Kcnj8», «Rgs5»	Pericytes	<ul style="list-style-type: none"> Ando et al. (2022) DOI:https://doi.org/10.1016/j.devcel.2022.04.019
1	«Klk6»	WM Oligodendrocytes	<ul style="list-style-type: none"> Elisa M. Floriddia <i>et al.</i> (2020) doi.org/10.1038/s41467-020-19453-x 	10	«Gpr17»	Differentiating OPCs	<ul style="list-style-type: none"> Robert Marisca et al. (2020) doi.org/10.1038/s41593-019-0581-2
2	«Cldn5» «Ly6c1»	Endothelial cells (hints to infection)	<ul style="list-style-type: none"> Jonas C. Schupp <i>et al.</i> (2021) doi: 10.1161/CIRCULATIONAHA.120.052318 Yuan-hui Li <i>et al.</i> (2022) doi: 10.3389/fimmu.2022.901672 	11	«Pdgfra»	Non-differentiating OPCs	<ul style="list-style-type: none"> Leanne E. Rivers et al. (2008) doi.org/10.1038/nn.2220
3	«Agt»	Astrocytes	<ul style="list-style-type: none"> Batiuk, M.Y. <i>et al.</i> (2020) doi.org/10.1038/s41467-019-14198-8 Jordan L Pauli <i>et al.</i> (2022) doi.org/10.7554/eLife.81868 	12	«Prph» «Fam19a4»	Neurons (Sensory)	<ul style="list-style-type: none"> Hsiu-Chan Lin <i>et al.</i> (2021) DOI:https://doi.org/10.1016/j.stemcr.2021.07.006 Dylan C. Sarver <i>et al.</i> (2021) doi: 10.1021/acschemneuro.0c00757
4	«Ptgds»	GM Oligodendrocytes	<ul style="list-style-type: none"> Elisa M. Floriddia <i>et al.</i> (2020) doi.org/10.1038/s41467-020-19453-x 	13	«Sfrp5»	Unknown (maybe adipocytes)	<ul style="list-style-type: none"> Rui He <i>et al.</i> (2024) DOI: 10.1042/CS20240179
5	«Tac1»	Glutaminergic neurons	<ul style="list-style-type: none"> Laura E. Mickelsen <i>et al.</i> (2019) doi.org/10.1038/s41593-019-0349-8 	14	«Sec14l5»	Oligodendrocytes	<ul style="list-style-type: none"> Noriyuki Ouchi <i>et al.</i> (2010) doi: 10.1126/science.1188280
6	«Mrgprd»	Mrgprd+ DRG neurons (innervate polymodal nociceptors)	<ul style="list-style-type: none"> Kaikai Wang <i>et al.</i> (2021) doi.org/10.1038/s41422-021-00479-9 Fu-Chia Yang <i>et al.</i> (2013) doi: 10.1016/j.celrep.2013.11.005 	15	«Tmem212»	Ependymal cells	<ul style="list-style-type: none"> Hussein Ghazale <i>et al.</i> (2019) doi: 10.1016/j.stemcr.2019.04.001
7	«Gira1» (``Fam19a1»)	Inhibitory interneurons	<ul style="list-style-type: none"> Lin Yao <i>et al.</i> (2020) doi.org/10.1016/j.ejphar.2021.174034 Dylan C. Sarver <i>et al.</i> (2021) doi: 10.1021/acschemneuro.0c00757 	16	«Ptgds»	GM Oligodendrocytes	<ul style="list-style-type: none"> Elisa M. Floriddia <i>et al.</i> (2020) doi.org/10.1038/s41467-020-19453-x
8	«Tmem119»	Microglial cells	<ul style="list-style-type: none"> Jordan L Pauli <i>et al.</i> (2022) doi.org/10.7554/eLife.81868 Quan Li <i>et al.</i> (2023) doi: 10.3389/fimmu.2023.1191188 	17	«Rgs5»	Pericytes	<ul style="list-style-type: none"> Ando et al. (2022) DOI:https://doi.org/10.1016/j.devcel.2022.04.019

Oligodendrocytes' subtypes (1)

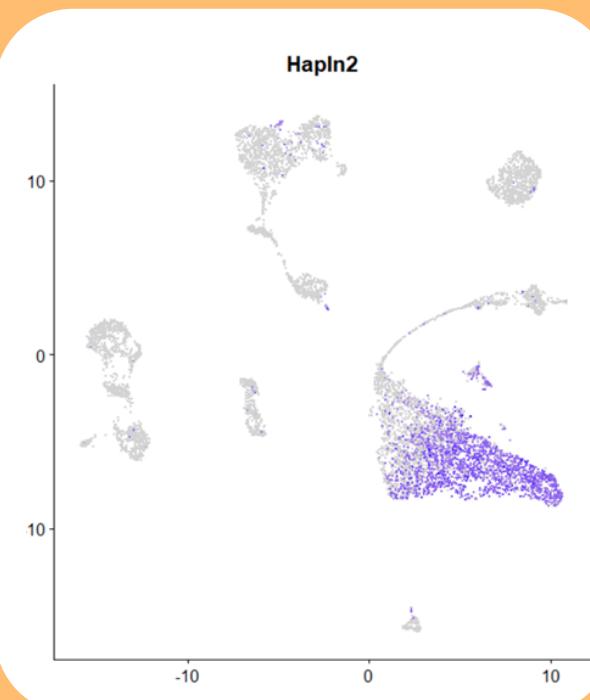
GENERAL MARKER*



MATURE OLS MARKER*



MYELINATION**

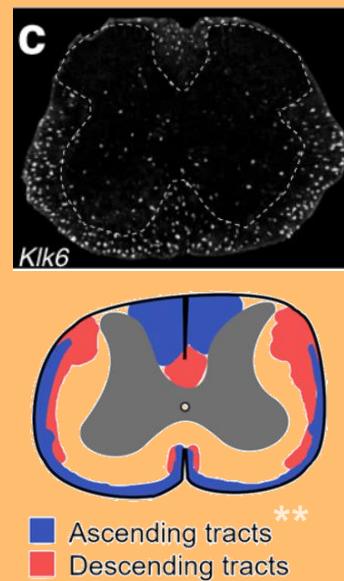
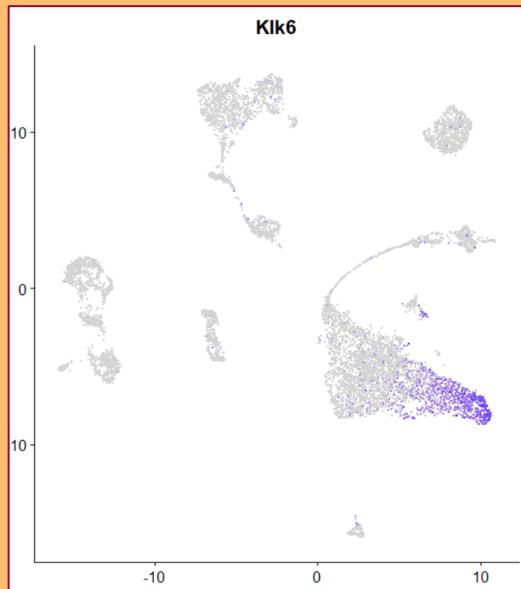


* Elisa M. Floriddia *et al.* (2020) doi.org/10.1038/s41467-020-19453-x

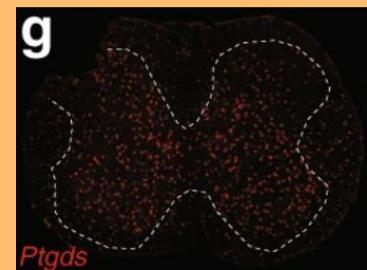
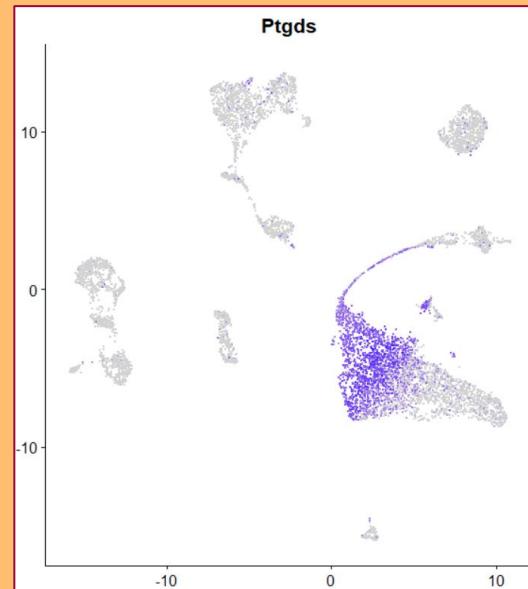
** Toshitaka Oohashi *et al.* (2002) doi.org/10.1006/mcne.2001.1061

Oligodendrocytes' subtypes (2)

WHITE MATTER OLs
POPULATION *



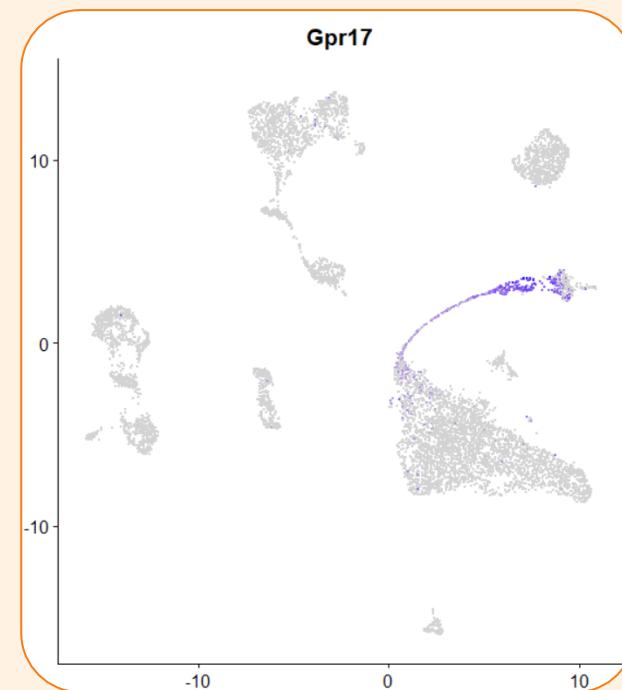
GREY MATTER OLs
POPULATION *



* Elisa M. Floriddia *et al.* (2020) doi.org/10.1038/s41467-020-19453-x

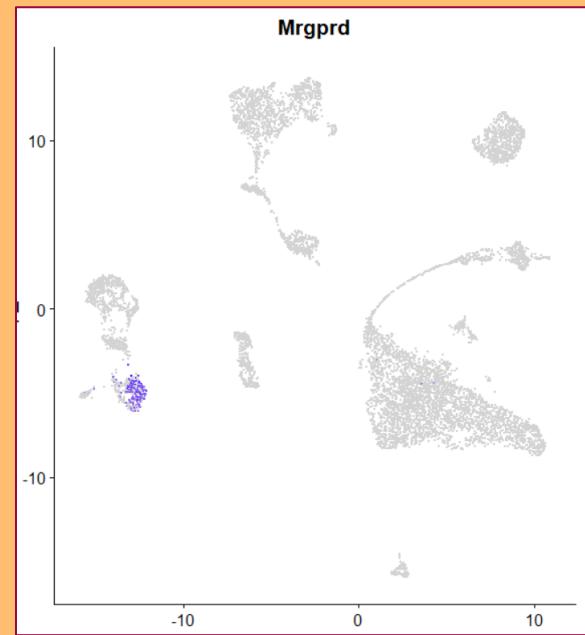
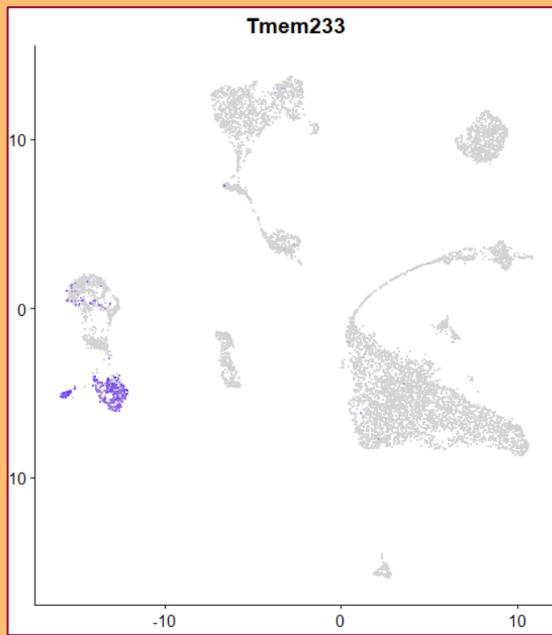
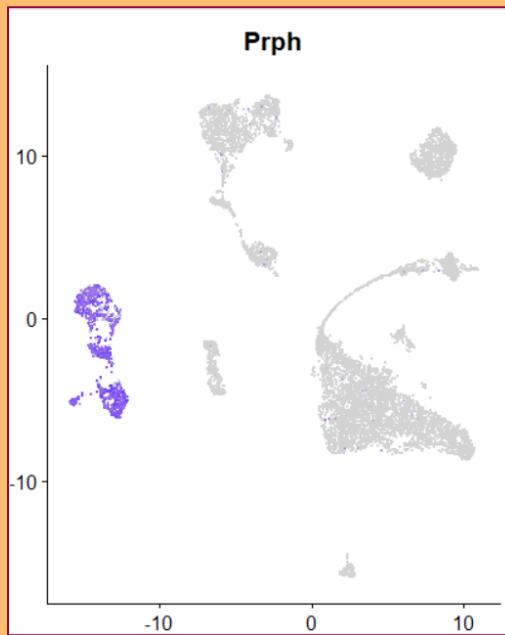
** https://www.researchgate.net/figure/Sketch-of-cross-sectional-views-of-a-rat-spinal-cord-a-The-white-matter-gray_fig2_344199287

Oligodendrocytes progenitor cells' subtypes



Pdgfra is downregulated when OLPs differentiate into oligodendrocytes, unlike lineage markers (Sox10). It's present only in undifferentiated cells (Leanne E. Rivers *et al.*, 2008), while Gpr17 is expressed by early differentiating OPCs (Robert Marisca *et al.*, 2020)

Cluster 6 neurons



Prph

PNS neurons & CNS regions with neural projections toward peripheral structures

Hsiu-Chuan Lin et al. (2021)
doi.org/10.1016/j.stemcr.2021.07.006

Tmem233

Trigeminal neurons. $\text{Log}_2\text{FC} = 5.57$

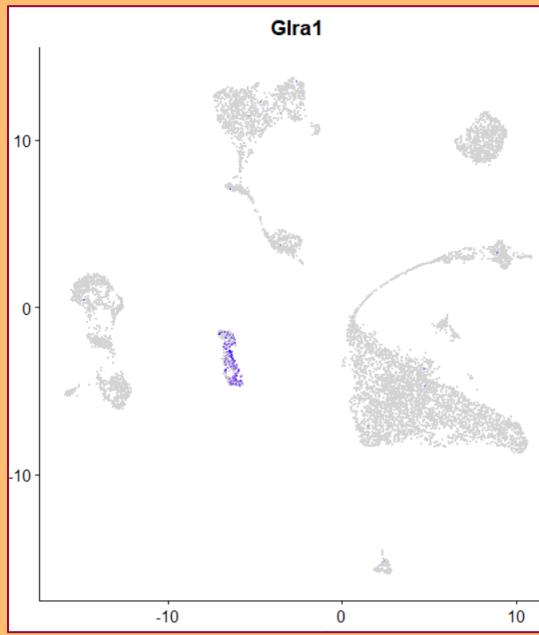
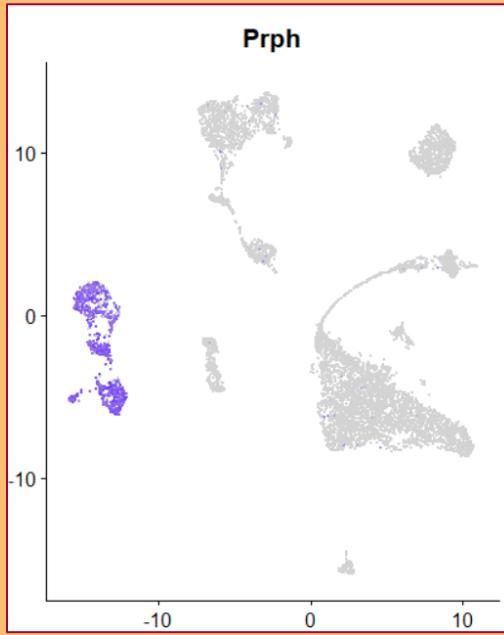
Sina Jami et al. (2023) doi.org/10.1038/s41467-023-37963-2

Mrgprd

Mrgrpd+ DRG neurons. $\text{Log}_2\text{FC} = 9.52$

Kaikai Wang et al. (2021) doi.org/10.1038/s41422-021-00479-9

Cluster 7 interneurons



Prph

PNS neurons & CNS regions with neural projections toward peripheral structures

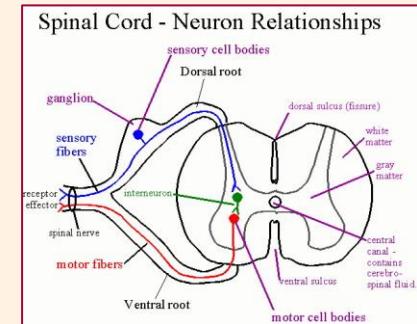
Hsiu-Chuan Lin et al. (2021)
doi.org/10.1016/j.stemcr.2021.07.006

Glra1

Inhibitory interneurons. Only part of CNS.

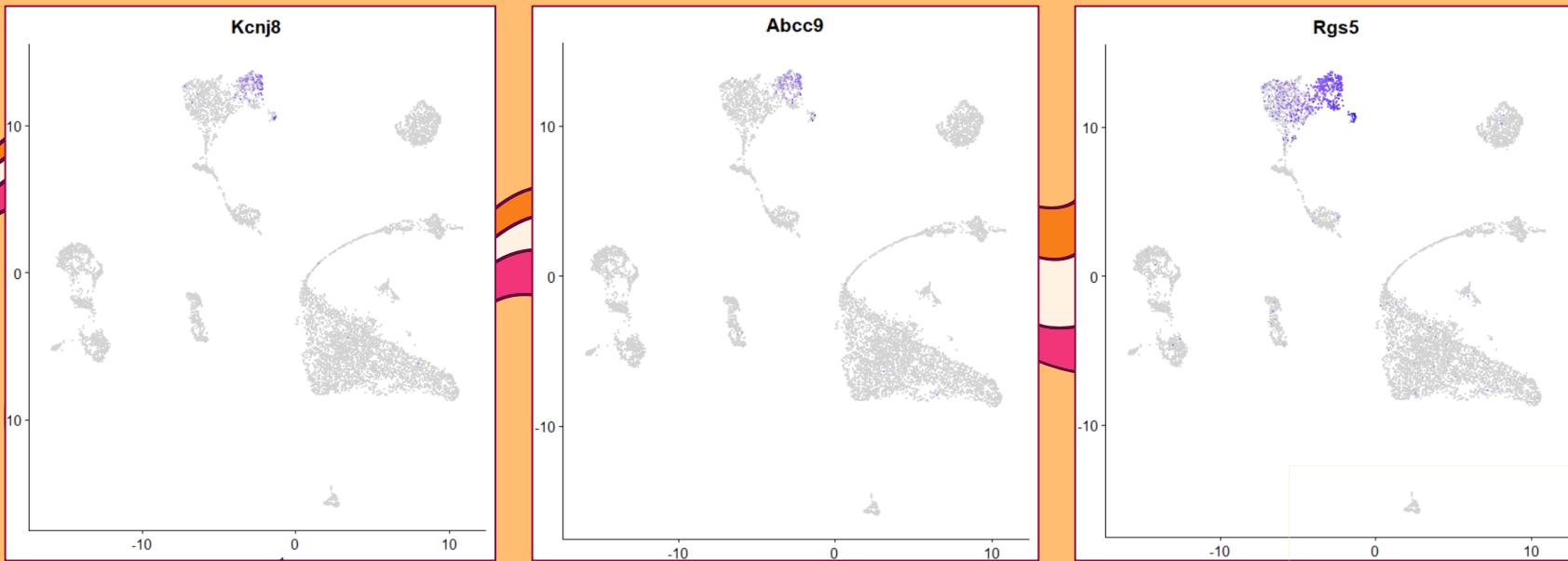
David I. Hughes and Andrew J. Todd (2020) doi:
10.1007/s13311-020-00936-0

- In the rat, inhibitory interneurons account for between 25 and 30% of all cells in laminae I and II, and approximately 40% of those in lamina III, with similar patterns being reported in the mouse. (David I. Hughes and Andrew J. Todd, 2020)
- The protein encoded by **Glra1** is a subunit of a pentameric **inhibitory glycine receptor**, which mediates **postsynaptic inhibition** in the **CNS**. (<https://www.ncbi.nlm.nih.gov/gene/2741>)
- Faint expression of **Fam19a1** further confirms it. (Dylan C. Sarver *et al.*, 2021)



Last consideration: is this mouse sick?

Kcnj8 and Abcc9 mRNA levels are low or undetectable in vascular smooth muscle cells (VSMC) but high in capillary pericytes (Koji Ando *et al.*, 2022)



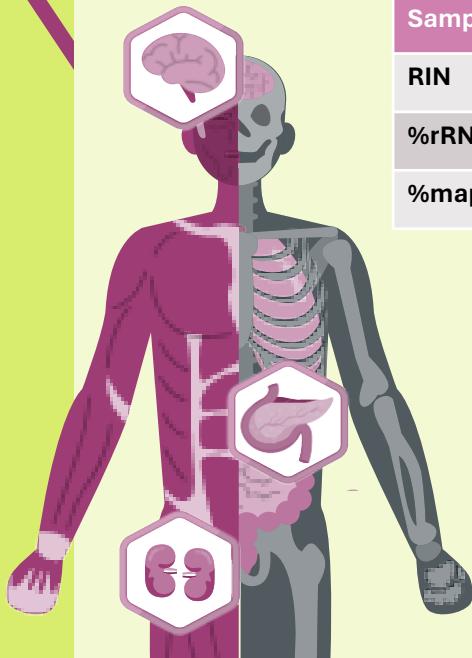
Among various molecules, RGS5 has been identified as signature marker of angiogenic tumor pericytes.
(Shayani Dasgupta *et al.*, 2021, doi.org/10.1038/s41418-021-00801-3)



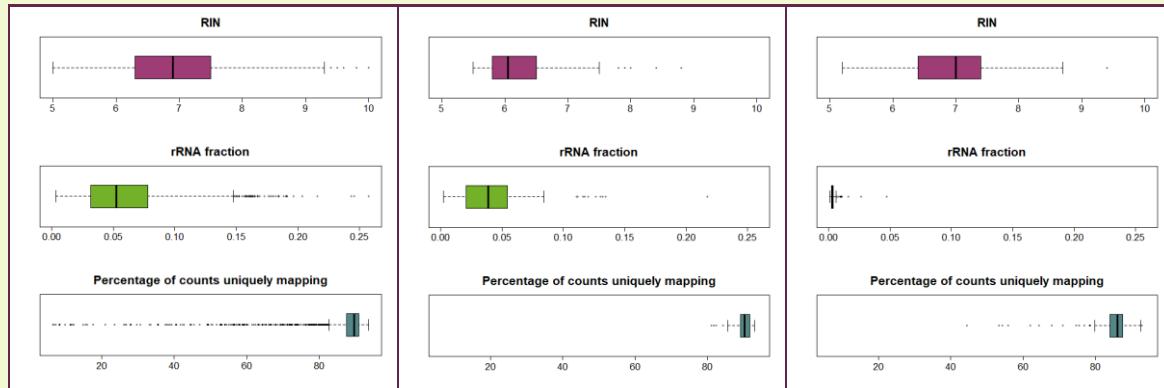
Bulk RNA-Seq analysis

BCG – Genomics and Transcriptomics course 2023/24
Flavia Leotta

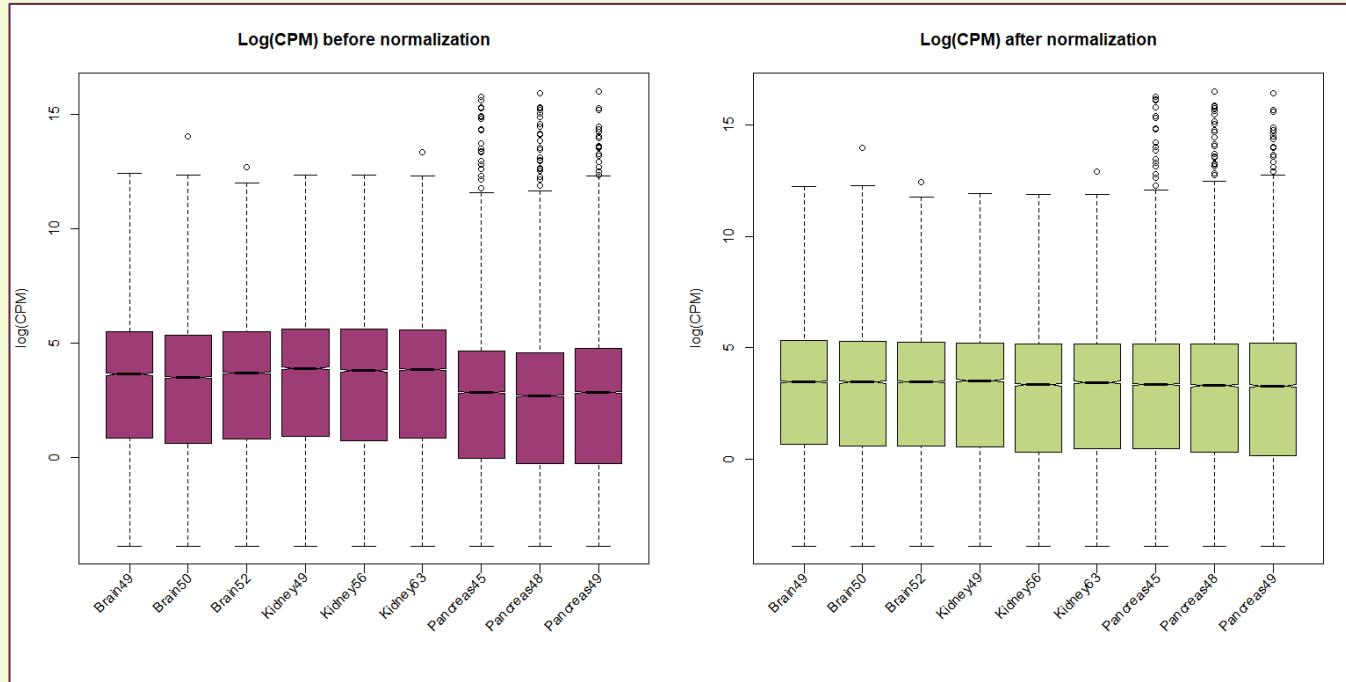
Quality control - 45866A



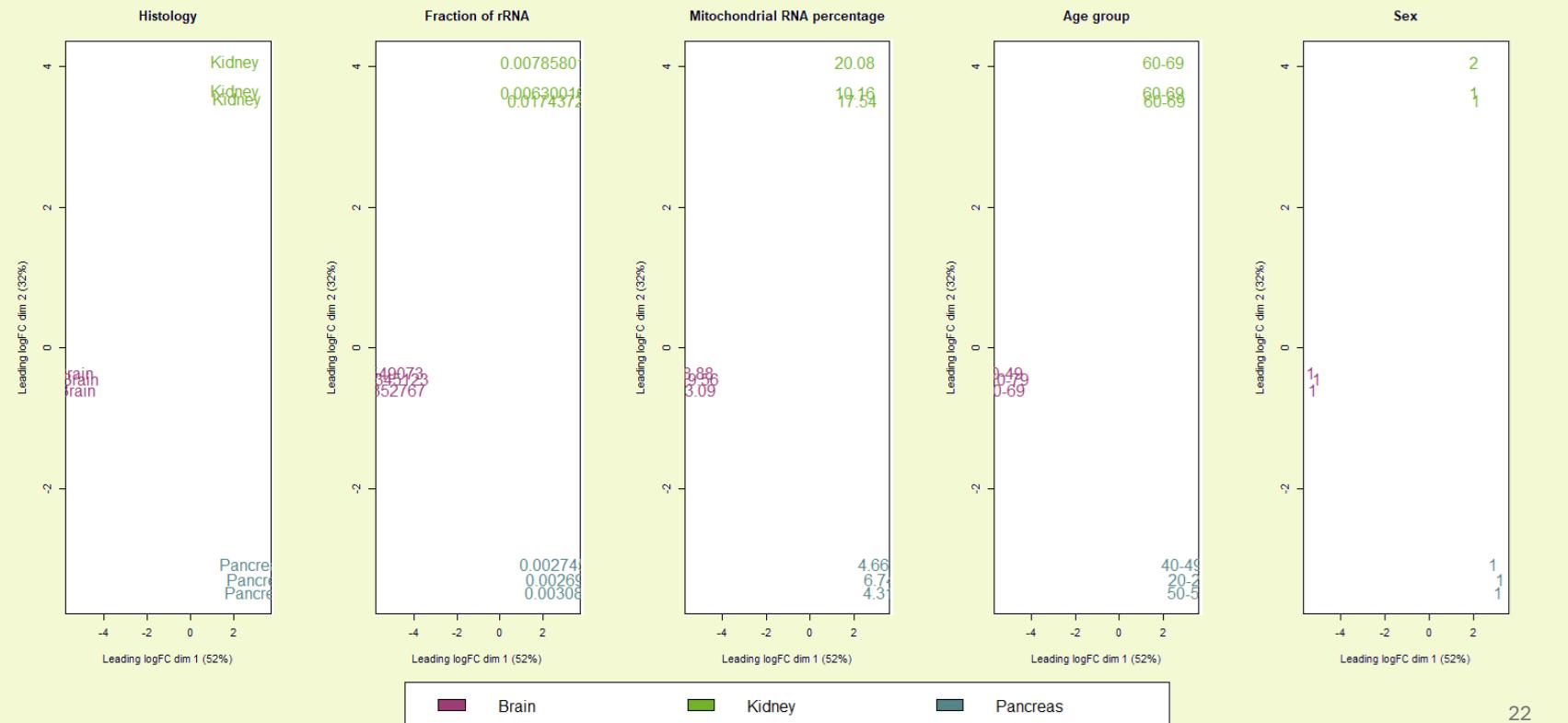
	B R A I N			K I D N E Y			P A N C R E A S		
Sample	49	50	52	49	56	63	45	48	49
RIN	6.8	6.7	7.4	7.3	7.8	8.0	7.6	7.5	7.1
%rRNA	3.5%	5.5%	3.4%	0.6%	1.7%	0.8%	0.3%	0.3%	0.3%
%map	91.5%	92.0%	88.9%	92.9%	89.7%	89.9%	86.0%	86.7%	86.5%



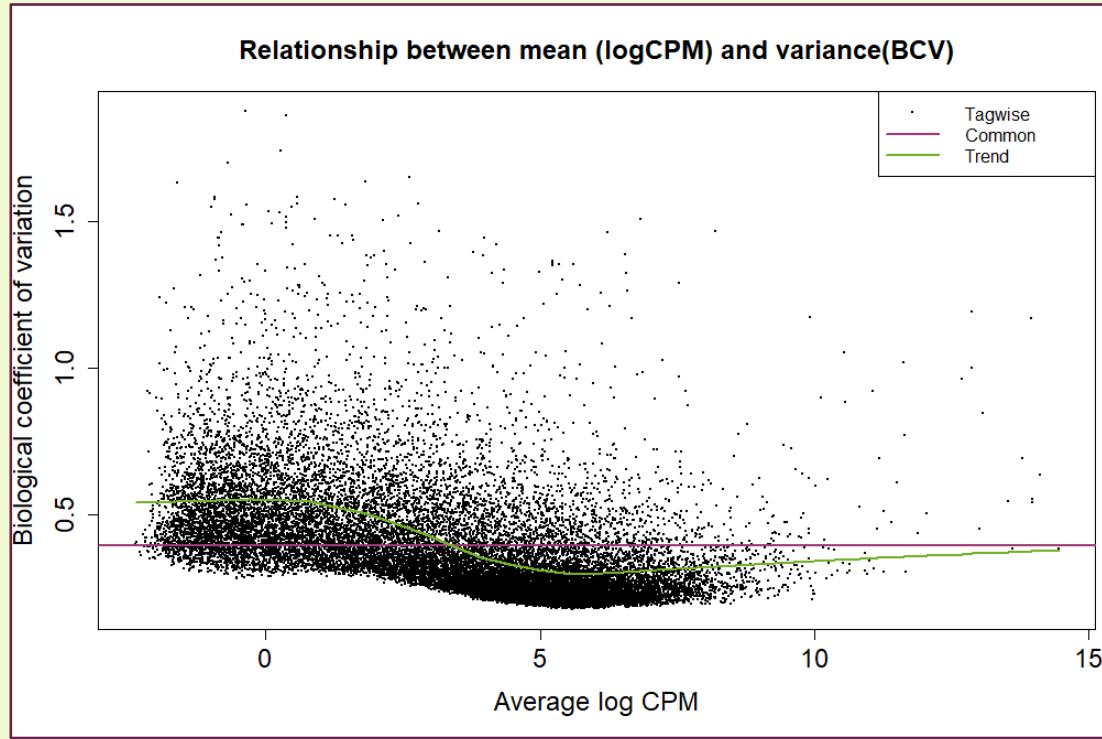
Normalization



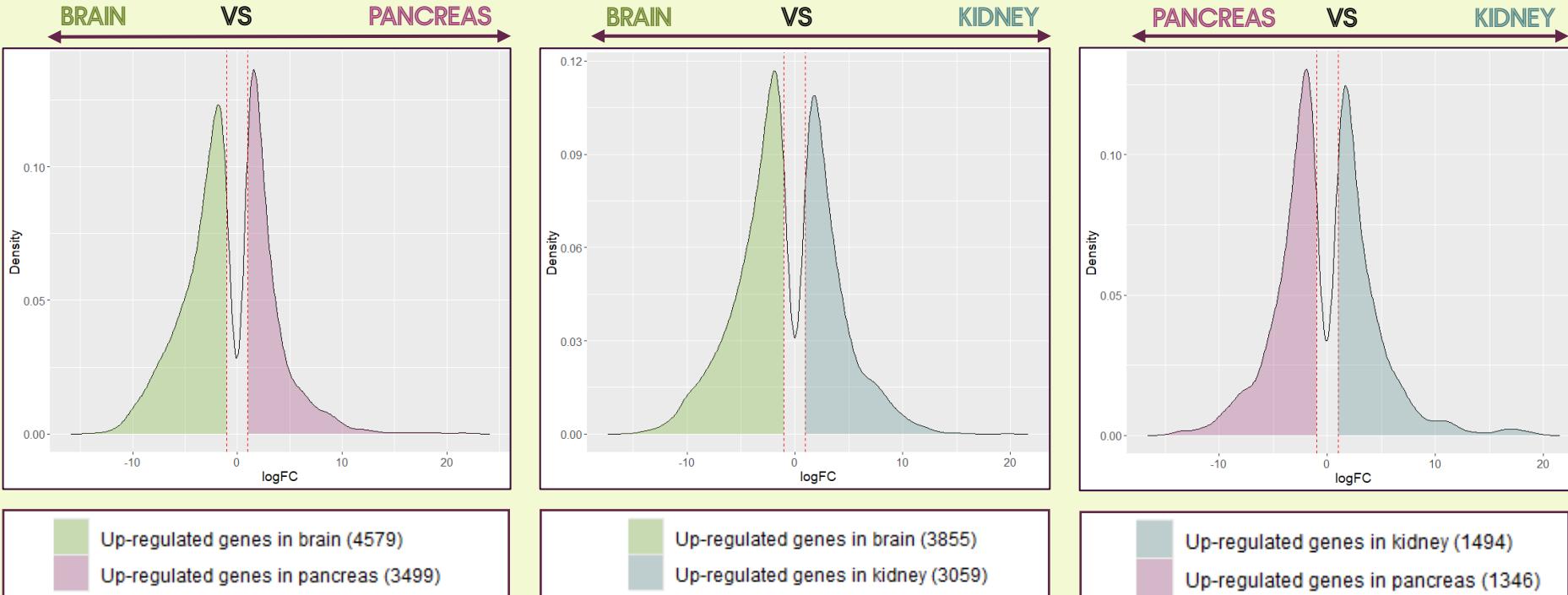
Samples 2D visualization



Mean and variance



DE genes - PAIRS (FDR < 0.01 and | LFC | > 1)



Up-regulated genes (FDR < 0.01 and | LFC | > 1)



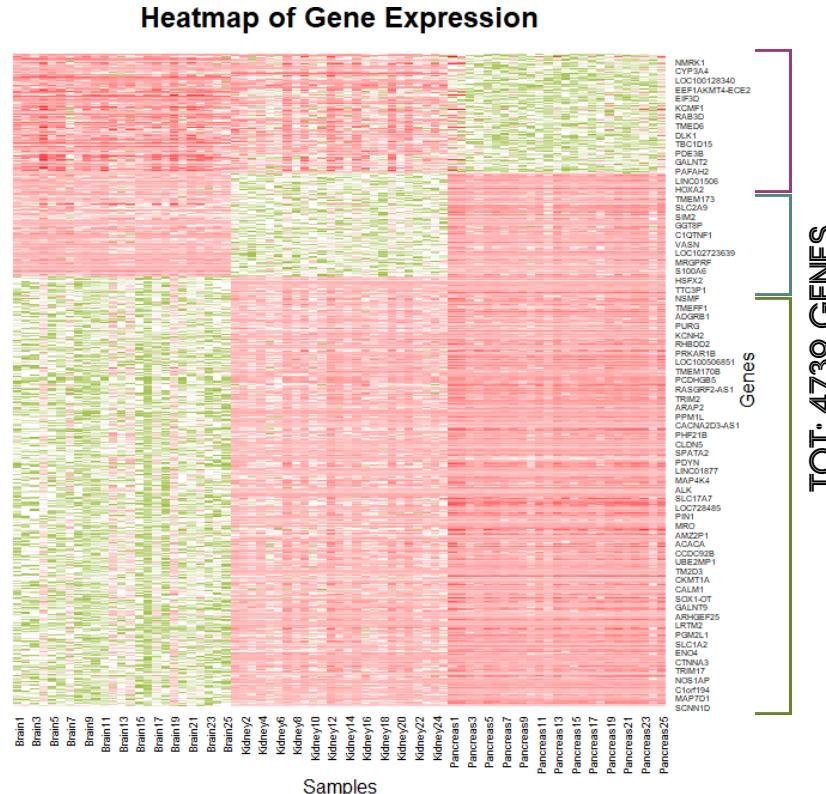
BRAIN
3119 GENES



PANCREAS
870 GENES



KIDNEY
750 GENES



Functional enrichment analysis over top 500 DE genes

On Gene Ontology Resource:

- Uploaded the top 500 DE genes in my three samples;
- Selected the biological process, molecular function and cellular component of the genes presenting the highest Fold Enrichment in comparison to the whole human genome.

On Enrichr (Human Gene Atlas):

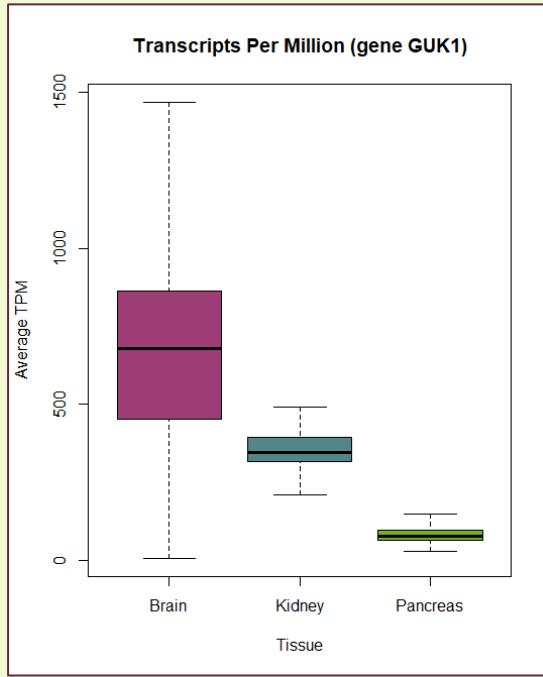
- Obtained ulterior proof that those genes are usually upregulated in the tissues we selected.
- The only difference is that for brain cells, the most probable cell type was «Prefrontal cortex», instead of «whole-brain» (that was the second one).
- Possible explanation: due to samples' position («Brain – cortex» (49), «Brain – Hippocampus» (50) and «Brain – Frontal Cortex (BA9)» (52)).

Functional enrichment analysis over top 500 DE genes

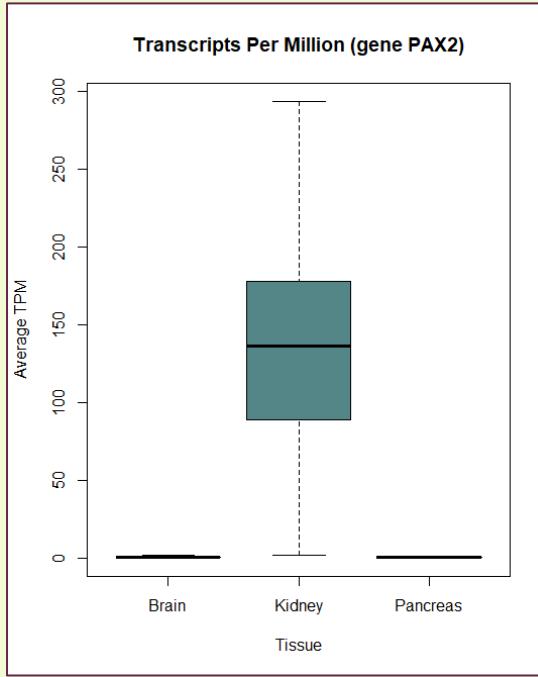
BRAIN TISSUE				KIDNEY TISSUE				PANCREAS TISSUE			
	GO annotation	Freq.	FDR	GO annotation	Freq.	FDR	GO annotation		Freq.	FDR	
BIOLOGICAL PROCESS	<i>purine deoxy-ribonucleoside diphosphate biosynthetic process</i>	2/2	3.67 E-02	<i>pronephric field specification</i>	2/2	1.79 E-02	<i>protein folding in endoplasmic reticulum</i>		5/11	2.46 E-03	
MOLECULAR FUNCTION	<i>G protein-coupled neurotransmitter receptor activity</i>	4/12	9.88 E-03	<i>polyspecific organization: proton antiporter activity</i>	2/2	3.32 E-02	<i>alpha-amylase activity</i>		5/5	1.97 E-05	
CELLULAR COMPONENT	<i>axon initial segment</i>	7/22	5.41 E-06	<i>integrin complex</i>	9/31	4.95 E-07	<i>methionine adenosyltransferase complex</i>		2/3	4.95 E-02	

Example - Comparison with all samples

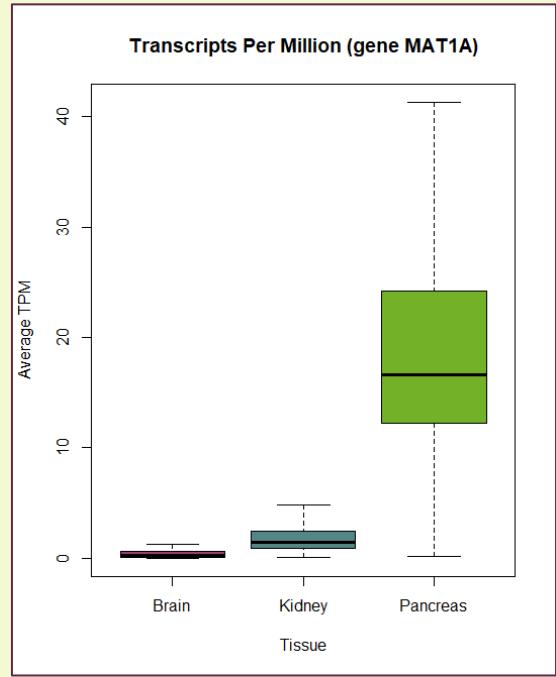
GUK1



PAX2



MAT1A



GO Annotations

	GUK1		PAX2		MAT1A	
NAME	Guanylate kinase		Paired box protein Pax-2		S-adenosylmethionine synthase isoform type-1	
MOLECULAR FUNCTION	ATP binding	GO:0005524	DNA binding	GO:0003677	ATP binding	GO:0005524
CELLULAR COMPONENT	Guanylate kinase activity	GO:0004385	DNA-binding transcription factor activity	GO:0003700	Identical protein binding	GO:0042802
	Protein binding	GO:0005515	Protein binding	GO:0005515	Protein binding	GO:0005515
	Kinase activity	GO:0016301	[6 more]		[2 more]	
	Cytosol	GO:0005829	Nucleus	GO:0005634	Cytosol	GO:0005829
Photoreceptor inner segment		GO:0001917	Microtubule organizing center	GO:0005815	Methionine adenosyltransferase complex	GO:0048269
[5 more]						