

*This document was created to introduce users to the Allen Brain Cell (ABC) Atlas and to provide an example use case of the tool and how to accomplish it. Last updated August 2025.*

## **Vignette Type: Experimental Design (w/Optional Coding)**

### **Specific Example: Whole Human Brain**

User:

Career: Undergraduate Students | Graduate Students | **Post-Docs** | **Senior Scientists/PI** | Teachers

Experience of Cell Types: Novice | Advanced Beginner | **Intermediate** | Expert

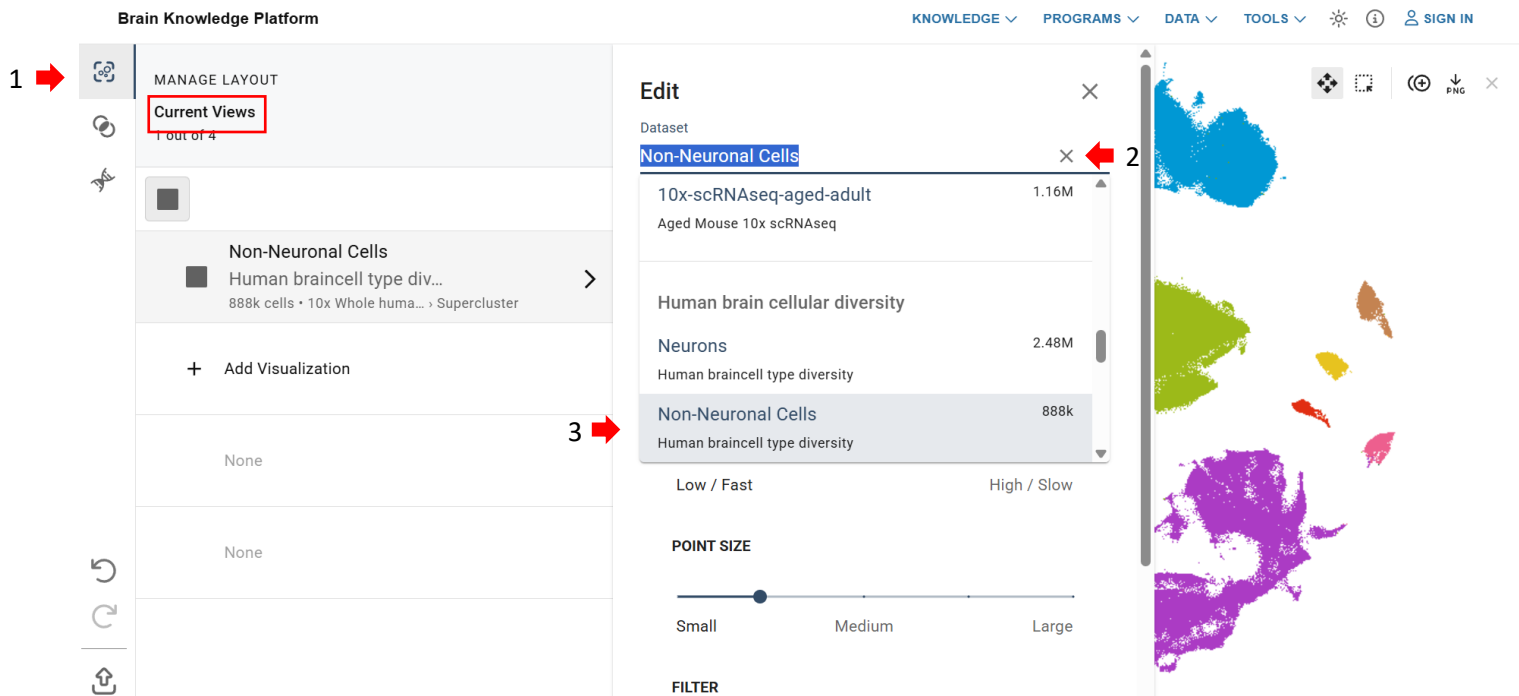
Research: Basic | **Translational**

Research Type: **Computational** | **Molecular** | Behavior

Experimental Model: Mouse | Rat | Non-Human Primate | **Human** | Invertebrate | Non-Traditional Vertebrate

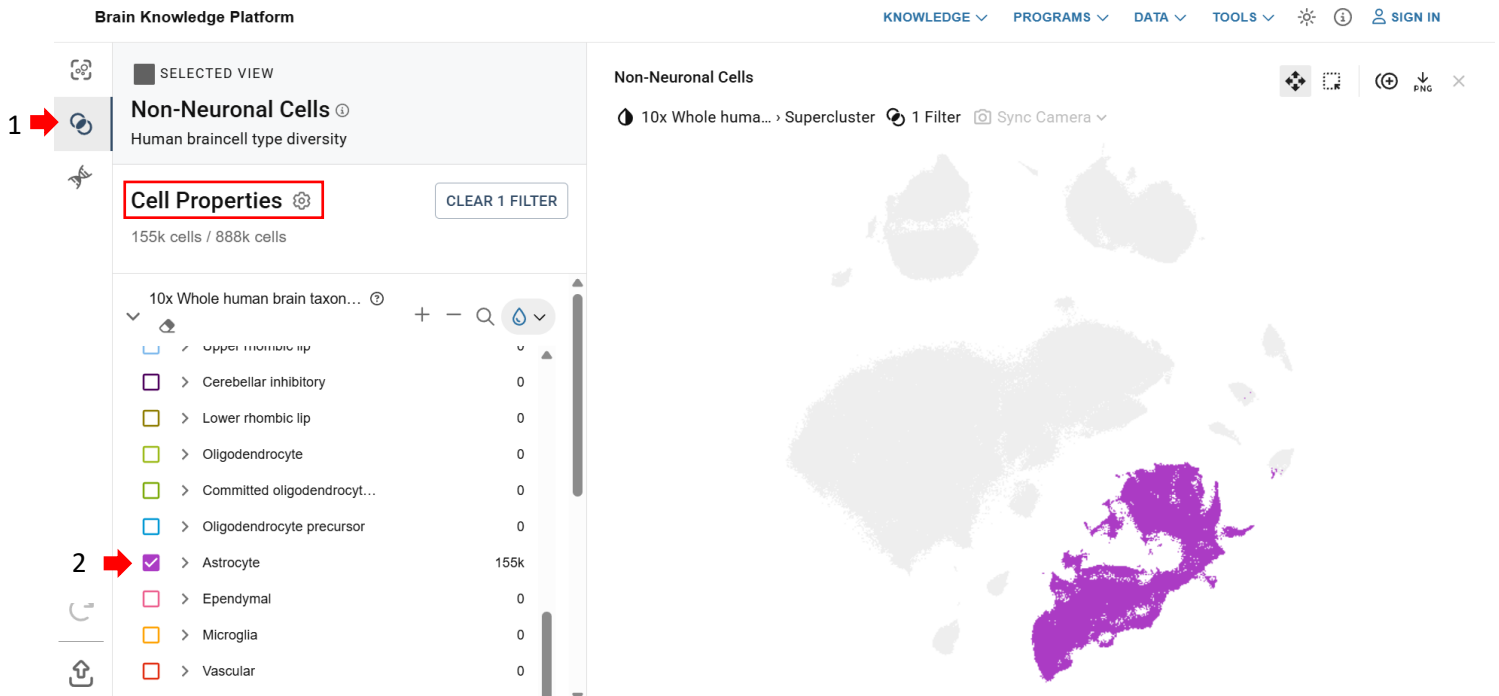
A researcher studies the role of astrocytes in the development of multiple sclerosis (M.S.), using postmortem spinal cord samples from humans with and without M.S. They are beginning to design studies for experiments in other commonly affected brain areas, such as the cerebellum. Therefore, they are interested in learning more about astrocyte diversity in other brain areas in the non-M.S. brain.

1. To view the non-neuronal nuclei in the dataset, the researcher clicks on the “Current Views” tab (step 1), clicks on the arrow next to “Dataset” (step 2) and clicks on “Non-Neuronal Cells” (step 3) under “Human brain cellular diversity”. [Link to view in ABC Atlas](#)

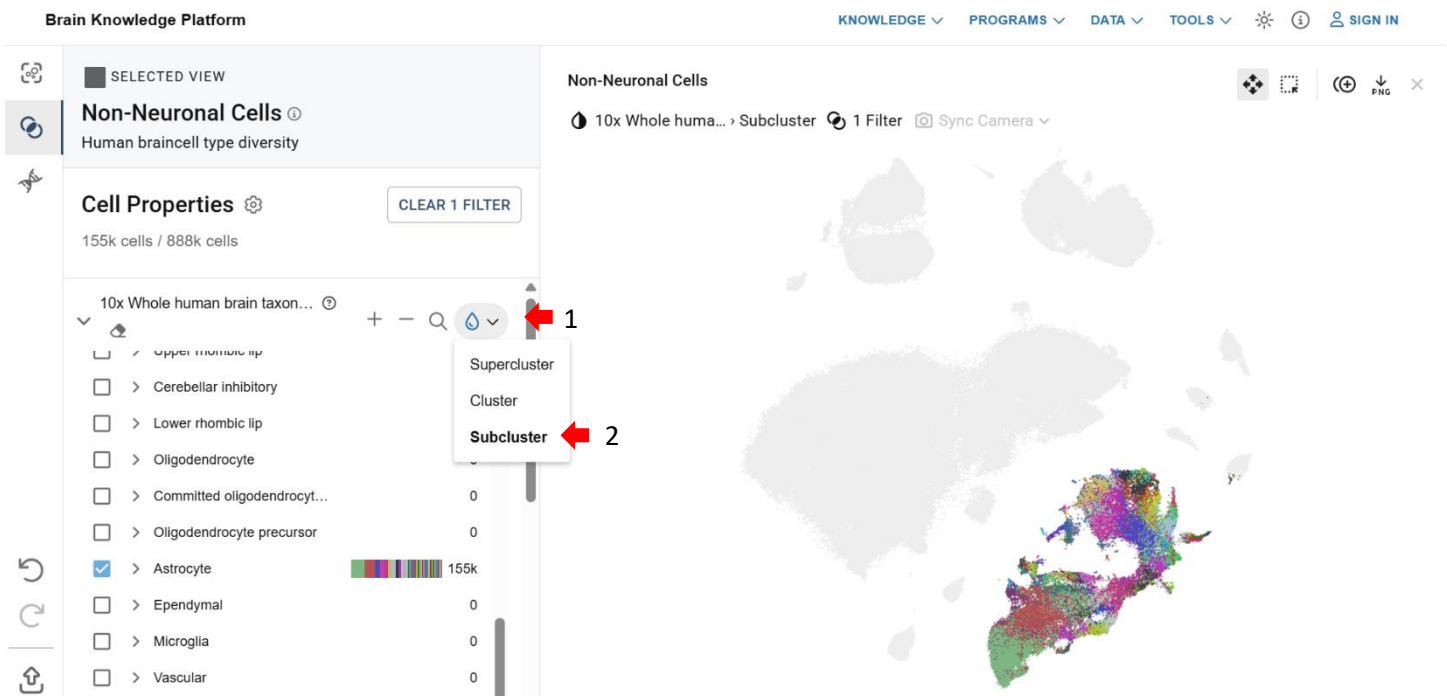


The screenshot displays the Brain Knowledge Platform interface. On the left, the 'MANAGE LAYOUT' sidebar shows the 'Current Views' tab selected (indicated by a red box and arrow 1). Below it, the 'Non-Neuronal Cells' dataset is listed with 888k cells. In the center, the 'Edit' panel shows the 'Non-Neuronal Cells' dataset selected (indicated by a red box and arrow 2). The 'Human brain cellular diversity' section is expanded, showing 'Non-Neuronal Cells' as the selected view (indicated by a red box and arrow 3). The right side of the interface shows a spatial map of the brain with colored regions representing different cell types.

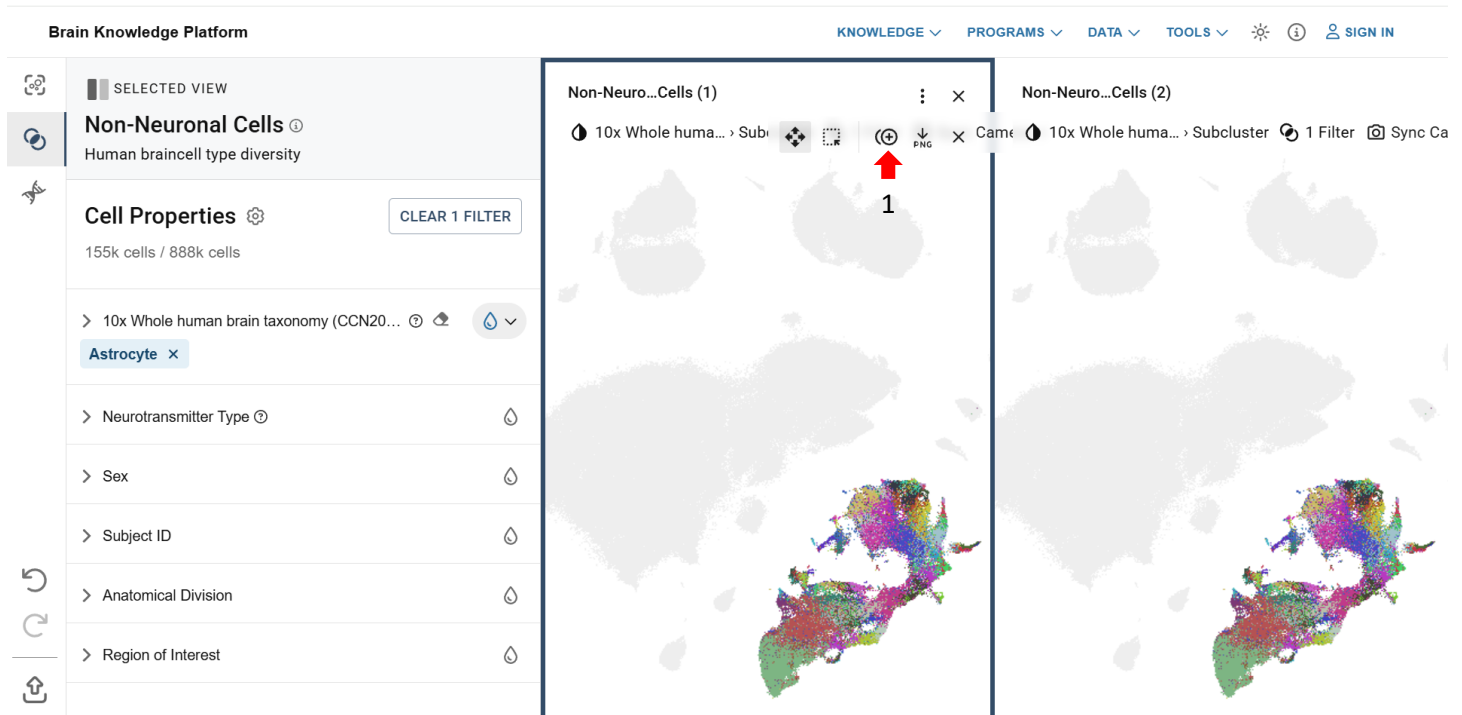
- To select only astrocytes, the researcher clicks on the “Cell Properties” tab (step 1), then clicks on the arrow next to “10x Whole human brain taxonomy” and checks the box next to “Astrocyte” (step 2). [Link to view in ABC Atlas](#)



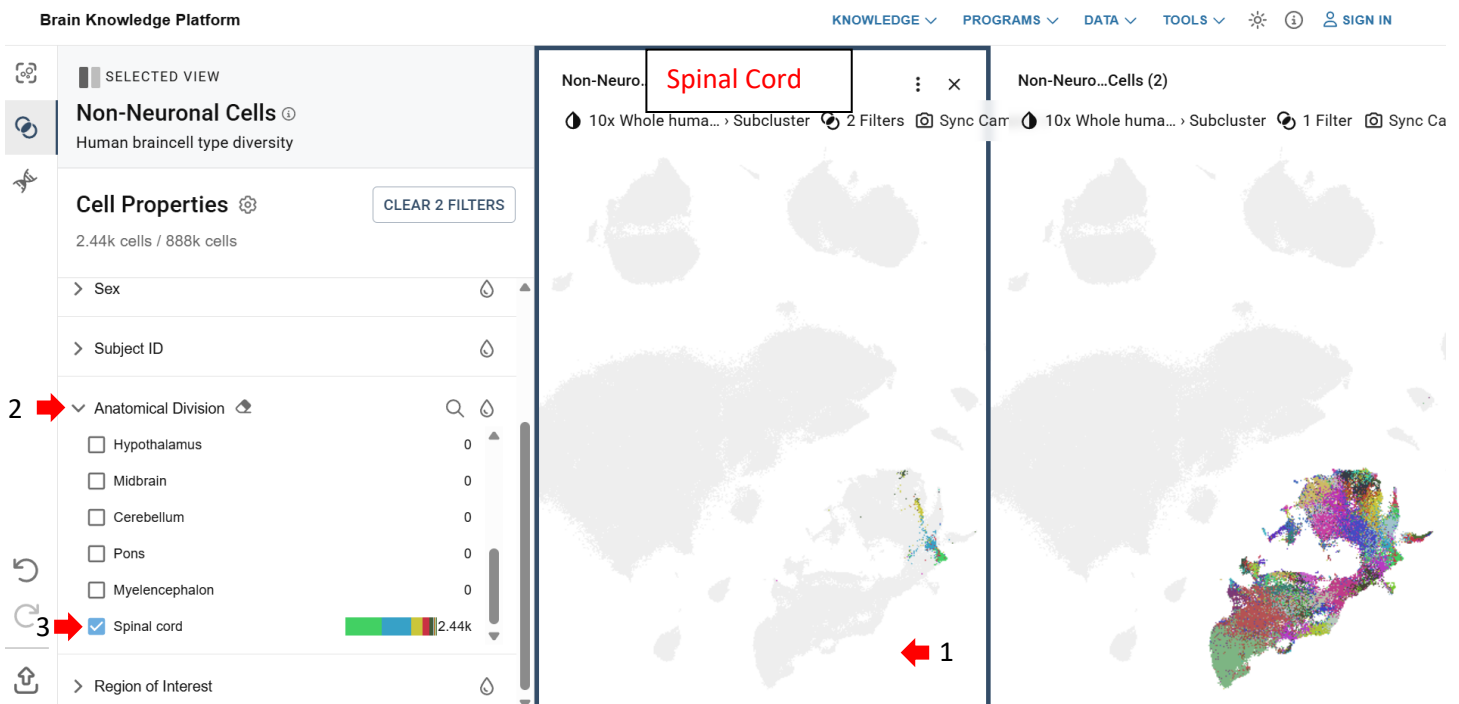
- To view the subclusters within the Astrocyte supercluster, the researcher clicks on the ink drop symbol (step 1) next to “10x Whole human brain taxonomy” and clicks on “Subcluster” on the pop-up menu (step 2). [Link to view in ABC Atlas](#)



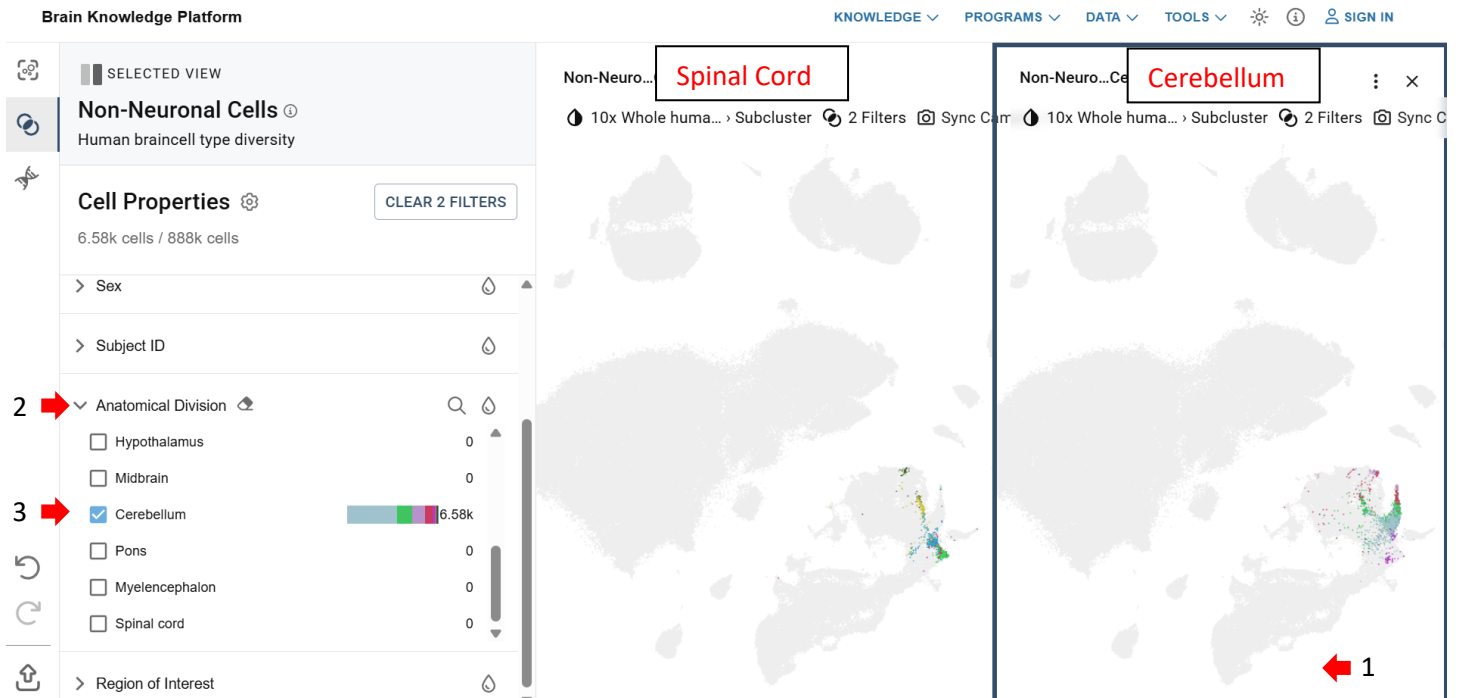
- To duplicate the view and look at two t-SNE plots at once, the researcher clicks the “Duplicate View” button (step 1). [Link to view in ABC Atlas](#)



- To look at just the spinal cord in the lefthand t-SNE, the researcher clicks on the lefthand t-SNE (step 1), clicks on the arrow next to “Anatomical Division” (step 2), and checks the box next to “Spinal cord” (step 3). [Link to view in ABC Atlas](#)



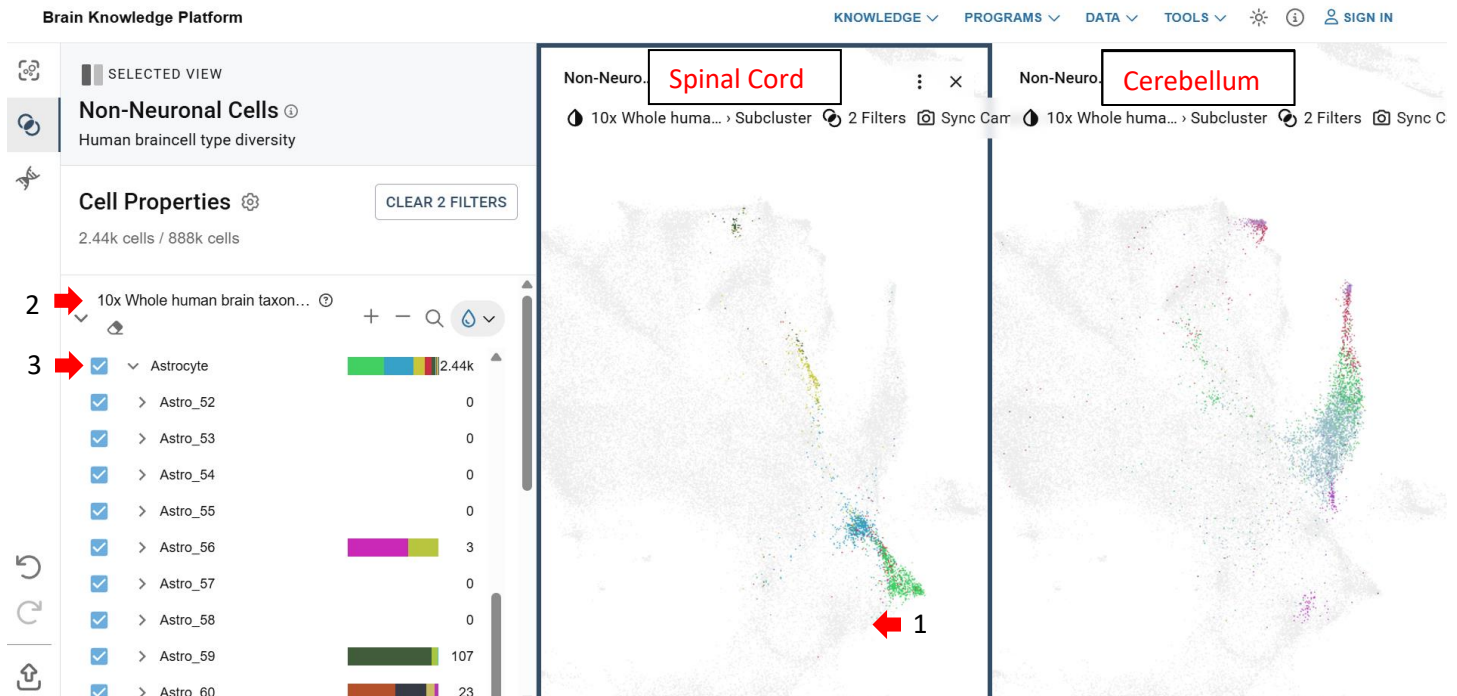
6. To look at just the cerebellum in the righthand t-SNE, the researcher clicks on the righthand t-SNE (step 1), clicks on the arrow next to “Anatomical Division” (step 2), and checks the box next to “Cerebellum” (step 3). [Link to view in ABC Atlas](#)



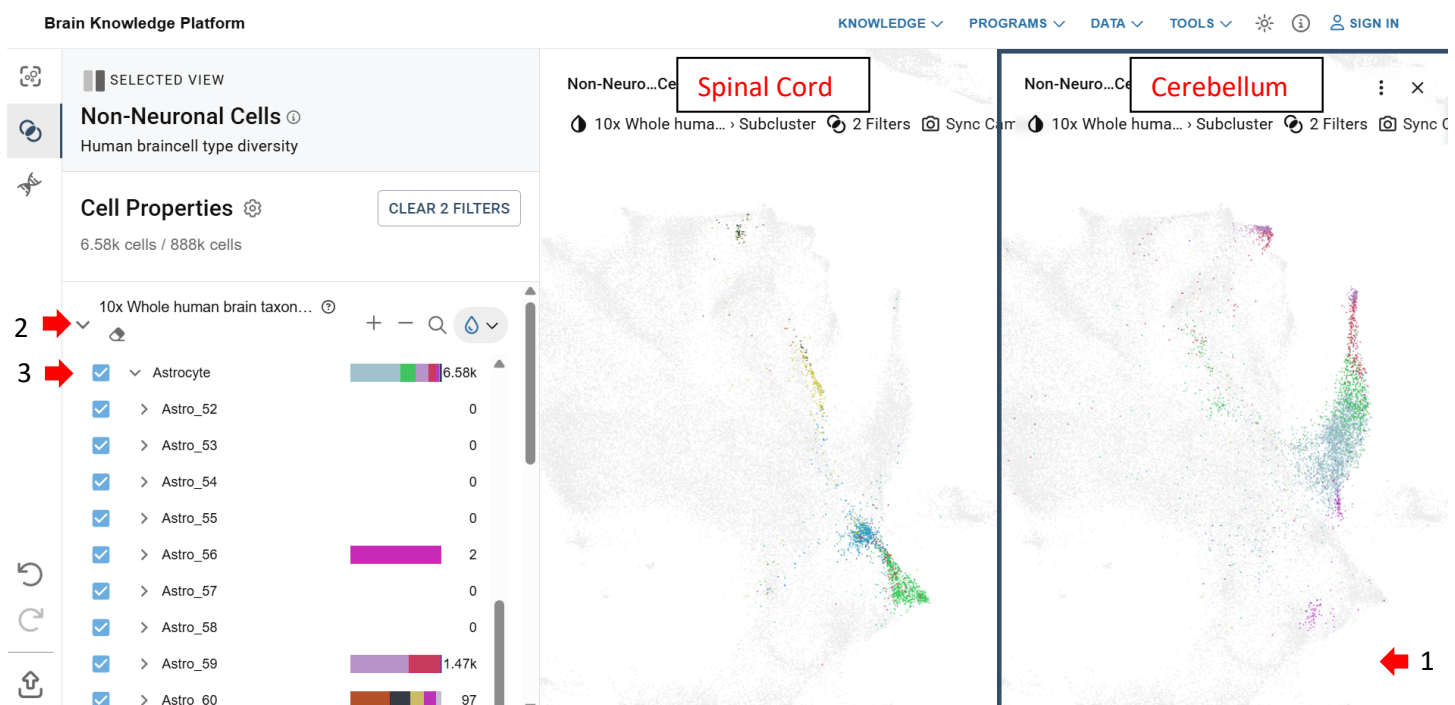
7. To zoom in on both t-SNE plots at the same time, the researcher clicks the “Sync Camera” button and checks the box to Non-Neuronal Cells (2) (step 1) and then zooms in by scrolling their mouse. [Link to view in ABC Atlas](#)



8. To view the clusters within the “Astrocyte” supercluster in the spinal cord, the researcher clicks on the lefthand t-SNE (step 1), clicks on the arrow next to “10x Whole human brain taxonomy” (step 2), then clicks on the arrow next to “Astrocyte” (step 3) to reveal the clusters. [Link to view in ABC Atlas](#)

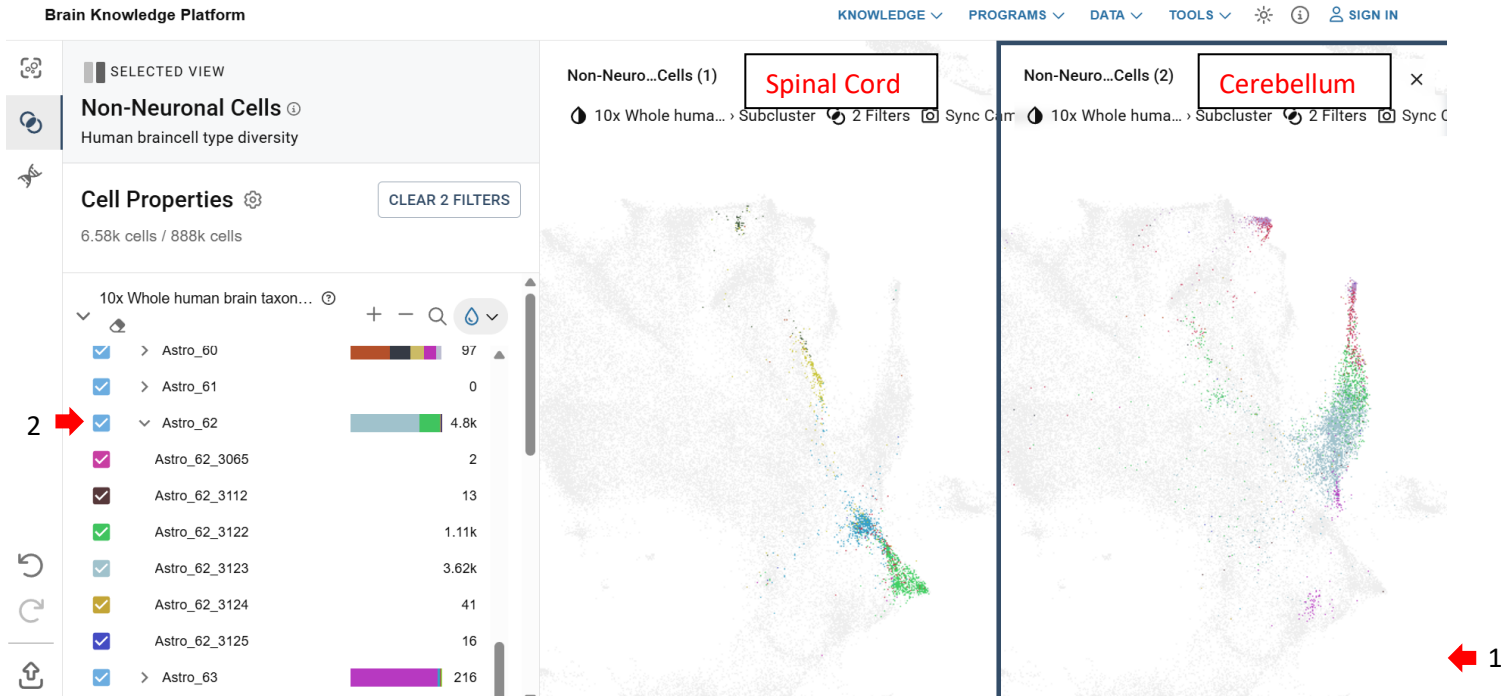


9. To view the clusters within the “Astrocyte” supercluster in the cerebellum, the researcher clicks on the righthand t-SNE (step 1), then click on the arrow next to “10x Whole human brain taxonomy” (step 2), then clicks on the arrow next to “Astrocyte” (step 3) to reveal the clusters. [Link to view in ABC Atlas](#)

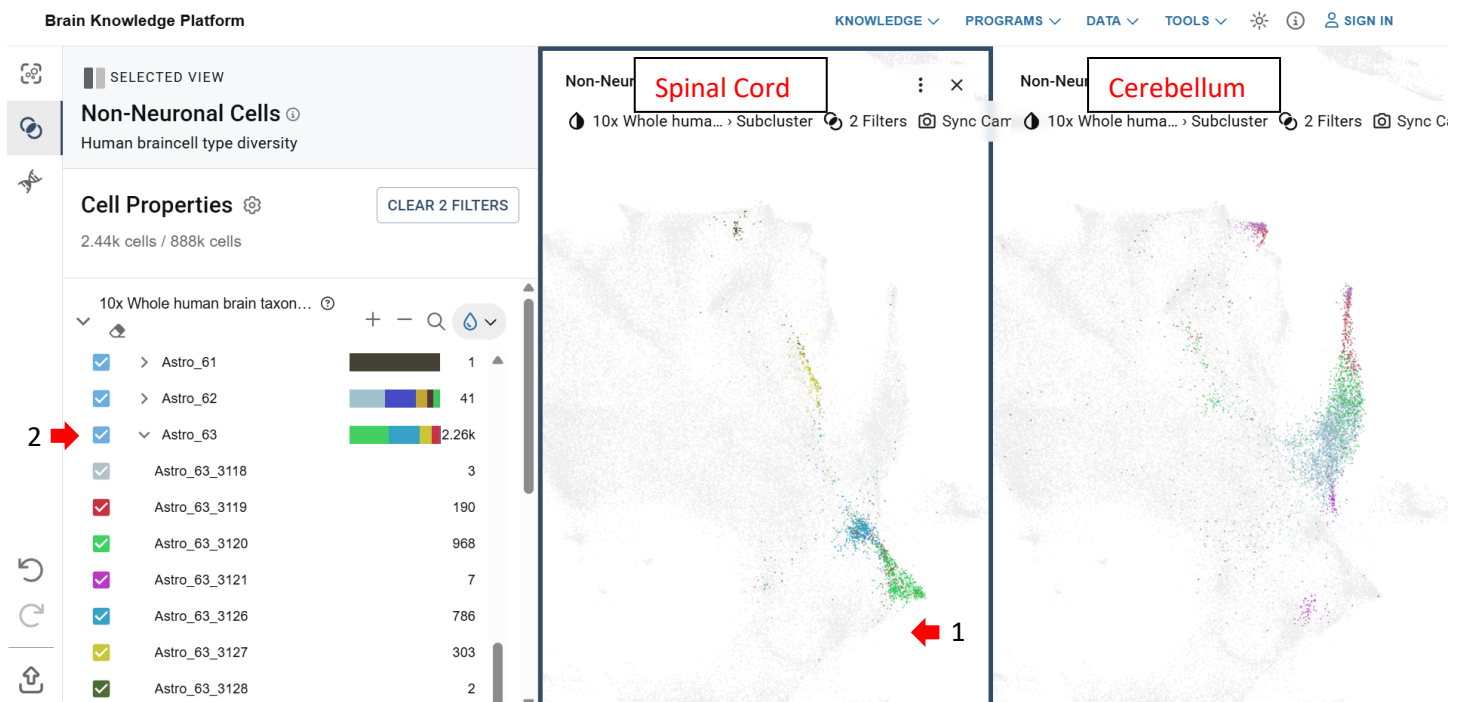




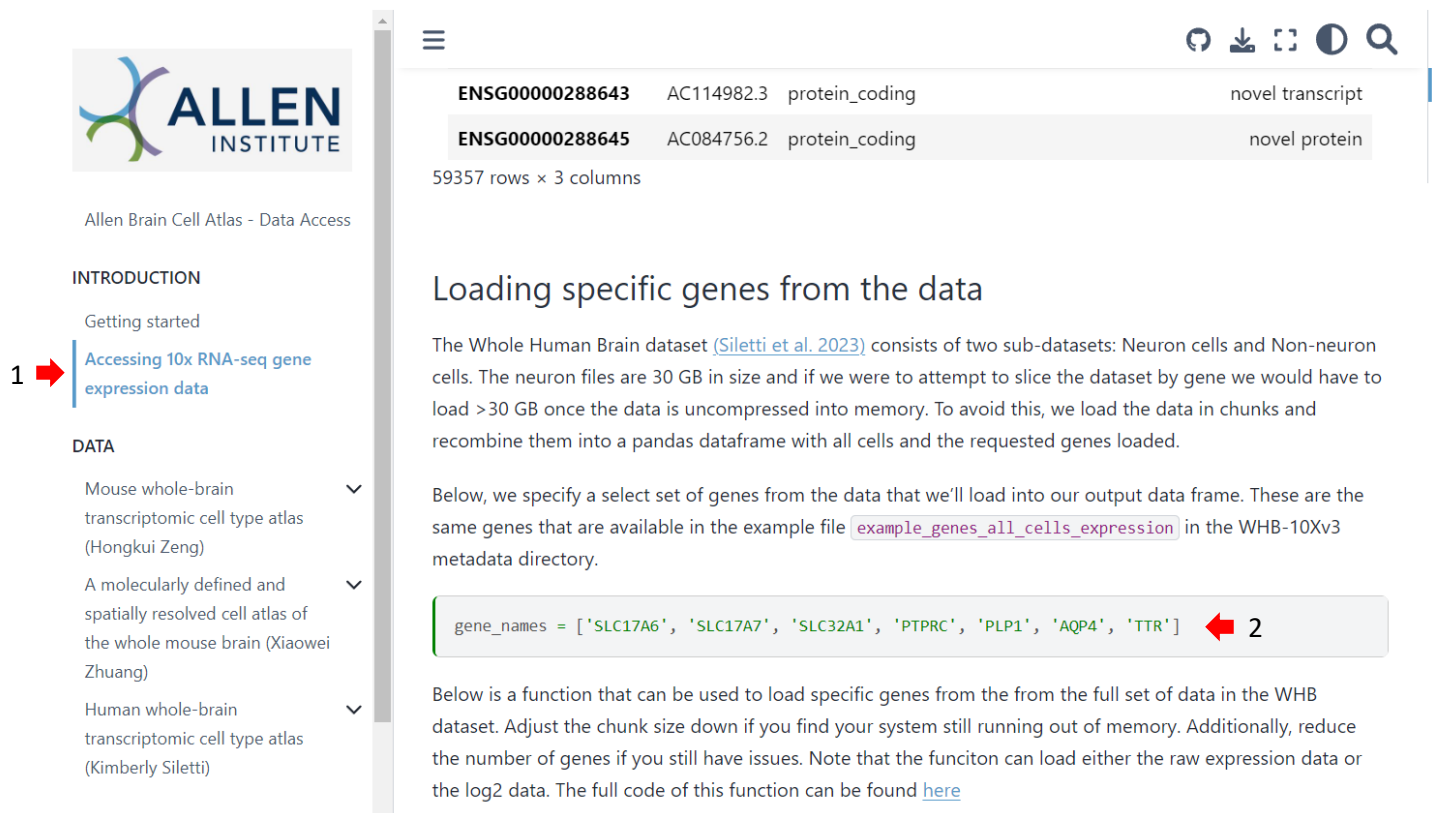
10. When looking at the cerebellum astrocyte clusters in step 9, the researcher notices that the majority of nuclei (4.8k) belong to the Astro\_62 cluster. To learn more about the Astro\_62 cluster, the researcher clicks on the righthand t-SNE (step 1), then clicks on the arrow next to Astro\_62 (step 2) to reveal the subclusters within the Astro\_62 cluster. Here the researcher sees that the majority of the nuclei (3.62k) belong to the Astro\_62\_3123 subcluster. [Link to view in ABC Atlas](#)



11. When looking at the spinal cord astrocyte clusters in step 8, the researcher notices that the majority of nuclei (2.26k) belong to the Astro\_63 cluster. To learn more about the Astro\_63 cluster, the researcher first clicks on the lefthand t-SNE (step 1), then clicks the arrow next to Astro\_63 (step 2) to reveal the subclusters within the Astro\_63 cluster. Here the researcher sees that the majority of the nuclei (968) belong to the Astro\_63\_3120 subcluster. [Link to view in ABC Atlas](#)



12. To learn more about the subcluster Astro\_63\_3120 (found in the spinal cord) and subcluster Astro\_62\_3123 (found in the cerebellum), the researcher goes to the [subcluster annotation table](#) to find the marker genes for these subclusters. In the table, the researcher finds the marker genes 'LINC01094' 'CD44' 'GFAP' 'ID3' 'APLNR' 'AL096709.1' 'AL627316.1' 'HSPB8' 'AC012405.1' 'FOS' for Astro\_63\_3120 and marker genes 'CPAMD8' 'GFAP' 'AC012405.1' 'AL627316.1' 'AC073941.1' 'CD44' 'AC097450.1' 'SLC14A1' 'PAX3' 'TNC' for Astro\_62\_3123.
13. After finding marker genes from the subcluster annotation table, the researcher decides to compare the expression of these genes across the two subcluster populations. The researcher does this by accessing the [WHB Jupyter Notebooks](#), and running the “Getting started” notebook. Afterwards, they click on “Accessing 10x RNA-seq gene expression data” notebook (step 1), scroll to “loading specific genes from the data,” and replace the example gene names listed with the subcluster marker genes (step 2.) After running the “getting started” notebook, they run the cells under “loading specific genes from the data.” Since the researcher is looking at non-neuronal data, they only need to download the non-neuron dataset.



The screenshot shows the Allen Brain Cell Atlas - Data Access interface. On the left, a sidebar contains the Allen Institute logo and a navigation menu with sections: INTRODUCTION, DATA, and a list of notebooks. The 'Accessing 10x RNA-seq gene expression data' notebook is highlighted with a red arrow and the number 1. The main content area shows the notebook's title, a table of gene expression data, and a code cell for loading specific genes. The table has 59357 rows and 3 columns. The code cell shows a list of gene names: ['SLC17A6', 'SLC17A7', 'SLC32A1', 'PTPRC', 'PLP1', 'AQP4', 'TTR'] with a red arrow and the number 2 pointing to it.

Allen Brain Cell Atlas - Data Access

INTRODUCTION

Getting started

1 Accessing 10x RNA-seq gene expression data

DATA

Mouse whole-brain transcriptomic cell type atlas (Hongkui Zeng)

A molecularly defined and spatially resolved cell atlas of the whole mouse brain (Xiaowei Zhuang)

Human whole-brain transcriptomic cell type atlas (Kimberly Siletti)

ENSG00000288643 AC114982.3 protein\_coding novel transcript

ENSG00000288645 AC084756.2 protein\_coding novel protein

59357 rows × 3 columns

### Loading specific genes from the data

The Whole Human Brain dataset ([Siletti et al. 2023](#)) consists of two sub-datasets: Neuron cells and Non-neuron cells. The neuron files are 30 GB in size and if we were to attempt to slice the dataset by gene we would have to load >30 GB once the data is uncompressed into memory. To avoid this, we load the data in chunks and recombine them into a pandas dataframe with all cells and the requested genes loaded.

Below, we specify a select set of genes from the data that we'll load into our output data frame. These are the same genes that are available in the example file [example\\_genes\\_all\\_cells\\_expression](#) in the WHB-10Xv3 metadata directory.

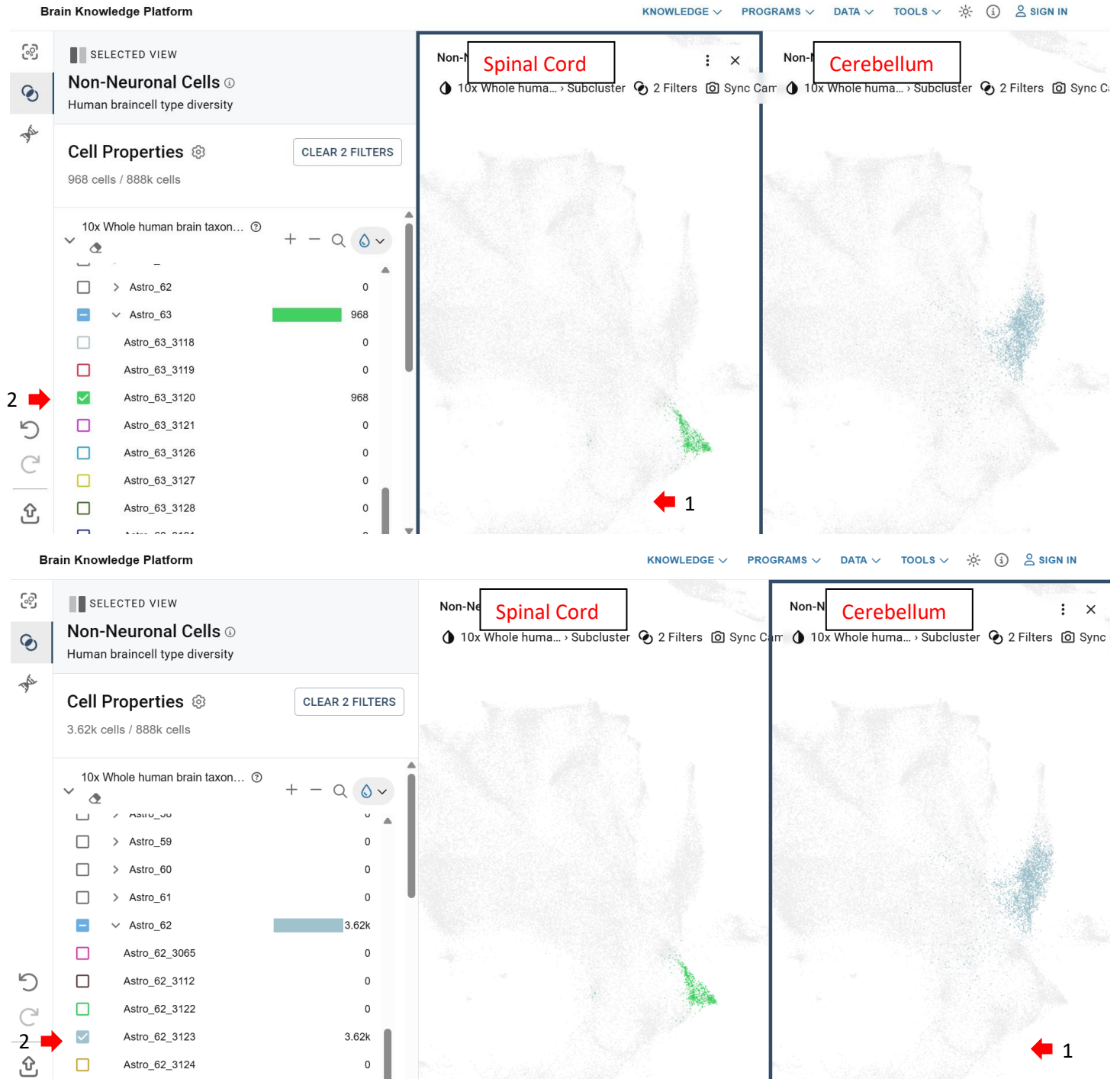
```
gene_names = ['SLC17A6', 'SLC17A7', 'SLC32A1', 'PTPRC', 'PLP1', 'AQP4', 'TTR']
```

2

Below is a function that can be used to load specific genes from the from the full set of data in the WHB dataset. Adjust the chunk size down if you find your system still running out of memory. Additionally, reduce the number of genes if you still have issues. Note that the function can load either the raw expression data or the log2 data. The full code of this function can be found [here](#)

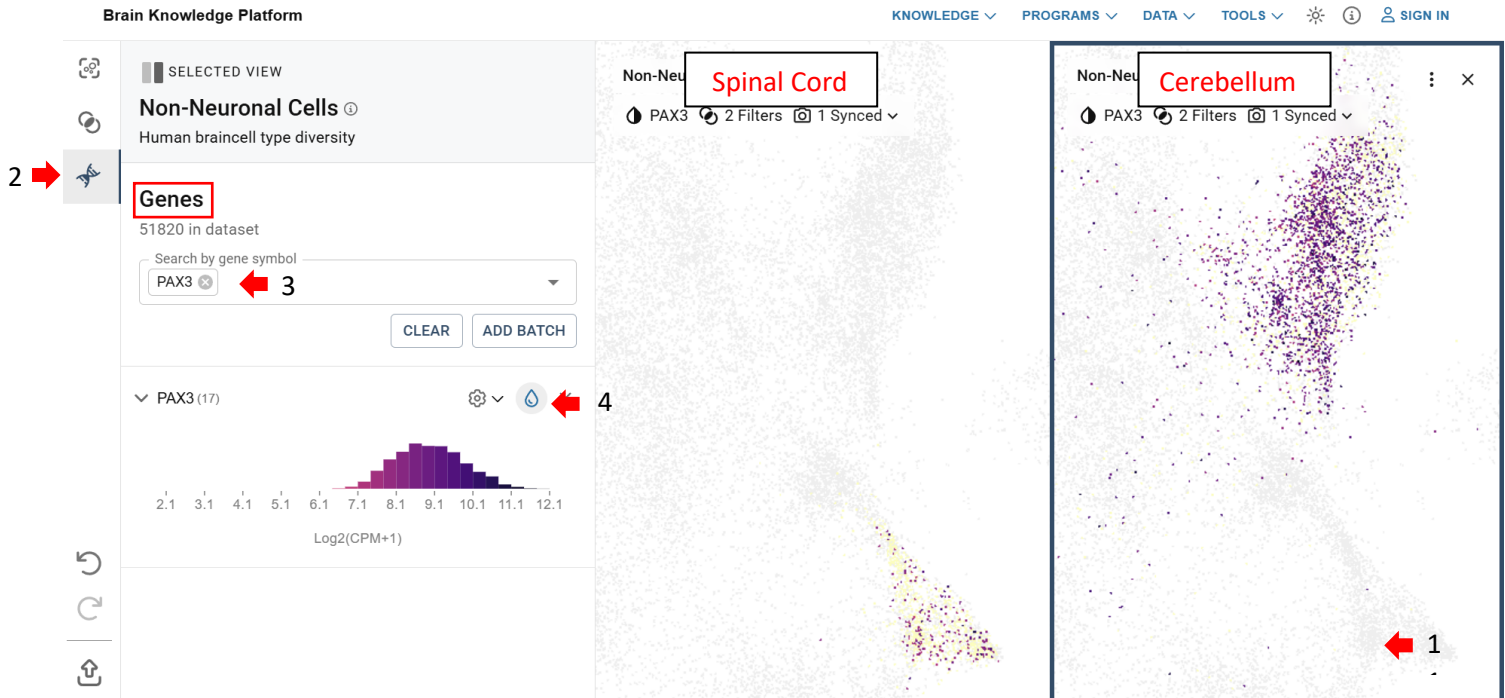
14. To compare gene expression across subclusters, the researcher clicks on the “Whole Human Brain 10x RNA-seq gene expression” tab under notebooks and clicks on “part 2.” Next, they change “example\_cells\_with\_genes” to “gene\_data” to use the genes they typed in for step 13, then run the cells.
15. After looking at the heatmaps of 'LINC01094' 'CD44' 'GFAP' 'ID3' 'APLNR' 'AL096709.1' 'AL627316.1' 'HSPB8' 'AC012405.1' 'FOS' and 'CPAMD8' 'GFAP' 'AC012405.1' 'AL627316.1' 'AC073941.1' 'CD44' 'AC097450.1' 'SLC14A1' 'PAX3' 'TNC', the researcher sees that PAX3 has higher expression in Astro\_62\_3123 than Astro\_63\_3120.

16. To isolate individual subclusters in the t-SNEs, the researcher then filters the lefthand spinal cord t-SNE by the Astro\_63\_3120 subcluster and the righthand cerebellum t-SNE by the Astro\_62\_3123 subcluster by clicking on the respective t-SNE plots (step 1) and then checking only those boxes under the “10x Whole human brain taxonomy” tab (step 2). [Link to view in ABC Atlas](#)





17. To color the Astro\_63\_3120 (spinal cord t-SNE) and Astro\_62\_3123 (cerebellum t-SNE) subclusters by the *PAX3* gene, the researcher clicks on a t-SNE plot (step 1), clicks on the “Genes” tab (step 2), types in “PAX3” (step 3), and clicks on the ink drop symbol (step 4), and then repeats these steps with the second t-SNE plot. [Link to view in ABC Atlas](#)



18. Now the researcher begins to design experiments based on their new knowledge of astrocyte diversity in the spinal cord vs. cerebellum.