Homework #3- Single-Cell RNA Sequencing Analysis

Dentate gyrus (DG) is part of the hippocampus involved in learning, episodic memory formation and spatial coding. The experiment from the developing DG comprises two time points (P12 and P35) measured using droplet-based scRNA-seq (10x Genomics Chromium).

In this homework, you're going to explore the data from this experiment.

Pre-task:

A. Install the requirements for **Tutorials 8 & 9**.
B. Load the dataset and save it to a variable by typing:

   adata = scv.datasets.dentategyrus()

C. This dataset is preprocessed, but was not normalized. Please normalize it by typing : **sc.pp.log1p(adata)**

Now when we're all set:

AnnData structure familiarity and visualization

1. Examine the object. How many genes and cells are included in the analysis? **(4 points)**
2. In this object, how many clustering methods are available? For each method, how many clusters were identified? **(6 points)**
3. Visualize the data using each clustering method. **(8 points)**

Differential Expression (DE)

4. Plot the top 20 differentially expressed genes **for each cluster** compared **to the rest**. **(8 points)**
5. What are the genes that best differentiate **Neuroblasts** from **the rest**? (Answer should be a data frame of the top 10 genes with their Log2FC scores and p-values, save it as a variable for GSEA questions). **(8 points)**
6. What are the genes that best differentiate **Granule immature** cells from **Granule mature** cells? (Same instructions as Q4) **(8 points)**
7. What are the genes that best differentiate **12-days** from **35-days**? (Same instructions as Q4) **(7 points)**

Gene Set Enrichment Analysis (GSEA)

8. Which pathways are enriched for **Neuroblasts** over **the rest**? **(8 points)**
9. Which pathways are enriched for **Granule Immature** cells over **Granule Mature cells**? **(8 points)**
10. Which pathways are enriched for **12-days** vs **35-days** cells? **(8 points)**

Pseudo-temporal analysis

11. What do you think is the least differentiated cell type in the lineage? (You can guess, you're not an expert) **(3 points)**
12. Define one of these cells as the root cell and do the pseudo-temporal analysis. What are the least differentiated and most differentiated cells in the lineage based on these results? **(8 points)**
13. Now do RNA velocity and latent time analyses. What are the results based on these analyses? **(8 points)**
14. Why did all these results differ from each other? What do we learn from each of them? **(8 points)**