Visium data analysis with Chipster 25.10.2022

chipster@csc.fi

PART I: Working with 1 slice

In this tutorial we analyze spatially resolved transcriptomics data generated with the Visium technology from 10x Genomics. The data contains two sagittal mouse brain slices generated using the Visium v1 chemistry.
We have downloaded the data (output files of the Visium Space Ranger pipeline) from here and imported in Chipster two sets of files for you: one set for an anterior section and one for a matched posterior section. We also imported a reference scRNA-seq dataset of ~14,000 adult mouse cortical cell taxonomy from the Allen Institute, generated with the SMART-Seq2 protocol.

Open Chipster: Go to **https://chipster.csc.fi/,** click on **Launch Chipster v4**, and log in.

1. Open training session

Click **Sessions** and select **course_spatially_resolved_transcriptomics** under **Training sessions**. Rename the session course_spatial_**your_first_name**

2. Find the right files and make a tar package of them

Open the tool manual for the tool **Spatial Transcriptomics / Seurat v4 -Setup and QC** (by clicking the **More info…** in the tool text). Check out the names of the input files needed. Select these 5 input files (=Space Ranger output files), and the tool **Utilites / Make a .tar package**. Name your package as mouse_anterior and run the tool.

3. Set up the Seurat object and perform quality control

Select the **mouse_anterior.tar** package generated in the previous step and the tool **Spatially resolved transcriptomics / Seurat v4 -Setup and QC.** Check the parameters, and set **Name of the sample** = **anterior1**. **Run** the tool.
Select the **QCplots.pdf** and click **Open in new tab**. Look at all the pages.

- What does a spot in the image represent now?
- Can you see any spatial variation in the number of transcripts?
- In which region do you see high mitochondrial transcript percentage? What can it mean?
- What is high hemoglobin percentage signaling?
- What would be the optimal limits for the mitochondrial transcript percentage (percent.mt) and hemoglobin transcript percentage (percent.hb)?

4. Filter spots, normalize data and detect highly variable genes with SCTransform

Select **seurat_spatial_setup_obj.Robj** and the tool **Seurat v4 – Filter spots, normalize with SCTransform and detect high-variance genes.** Would the default spot filtering parameters be good for this dataset, based on the QC plots? In order to follow the Seurat vignette, let's not remove any spots for now: set the mitochondrial and hemoglobin filtering parameters to **100**, and run the tool.

5. Gene expression visualization

Select **seurat_obj_sctransform.Robj** from the previous step and run the tool **Seurat v4 - Visualise gene expression** with the default parameters for genes Hpca and Ttr. Then run the tool again, but change **Minimum transparency = 0.1**.
Open the resulting **Feature_plot.pdf** files in new tabs and compare them.
- Are these two genes spatially variable based on these plots?
- What is the difference between these two plots?

6. Principal component analysis, clustering and visualisation

Select **seurat_obj_sctransform.Robj** from the step 4 and run the tool **Seurat v4 - PCA, clustering, and visualisation** using the default parameters. Open **UMAP_plot.pdf** in new tab.

- How many clusters are there? Does the coloring (= clustering) match the grouping found by UMAP?
- Are the clusters located in particular areas of the tissue?
- Cerebral cortex, also called gray matter, is the brain's outermost layer of nerve cell tissue. Which clusters correspond to the cortex?

7. Visualise clusters

Looking at the colorful spatial cluster plots can give you a migraine. To more easily detect where the clusters are located, let's visualize some of them one by one. Select **seurat_spatial_obj_pca_clust.Robj** from the previous step and the tool **Seurat v4 - Visualize clusters**. In the parameters, type **clusters to plot = 1, 2, 3, 4, 5, 6, 7, 8.** Open the resulting **spatialdimplot.pdf.**

- Which of the clusters corresponds to the outermost layer? Does it have any spots elsewhere in the tissue?
- Where are clusters 5 and 6 located?
- Can you spot the cortex clusters more easily now?

8. Identify spatially variable genes

Choose **seurat_spatial_obj_pca_clust.Robj** generated in step 6. Select tool **Seurat v4 - Identify spatially variable genes based on clusters.** Type clusters to compare to the parameter fields: **first cluster = 5** and **second cluster = 6**. Open the resulting **markerplot.pdf**, and compare it to the **spatialdimplot.pdf** for clusters 5 and 6.

-What are the 3 most spatially variable genes between these two clusters?

Open **spatially_variable_genes.tsv**.
-Which cluster is expressing which of these three genes? Are there differences in levels (see columns pct.1, pct.2 and avg_log2FC)?

Previous step was comparing two clusters. With the tool **Seurat v4 -Identify spatially variable genes using markvariogram** we can identify genes that have spatial patterning without taking cluster information or spatial annotation into account. This tool takes very long time to run, so we have run it for you. Open the file **Markvariogram.pdf**.

- What are the 6 most spatially variable genes?

9. Subset anatomical regions

Let's subset the frontal cortex spots. Choose **seurat_spatial_obj_pca_clust.Robj** generated in step 6 and tool **Seurat v4 - Subset out anatomical regions based on clusters.** Type **subset of clusters = 3,4,6,7** and run the tool. Open **subset.pdf**.

- Did you get the cortex spots subsetted?

10. Annotate spots by integrating spatial data with scRNA-seq data

Select **seurat_obj_subset.Robj** generated in the previous step and **allen_cortex.rds** reference file already uploaded in the session, and run the tool **Seurat v4 -Integration with single-cell data**. Open the resulting **reference_UMAP_plot.pdf**.

- What do you see in the plot? What does each dot represent?

Select **seurat_obj_integrated.Robj** and tool **Seurat v4 -Visualise integration**. In the parameters, type **Features to plot = L4**. Open **integration_plot.pdf**.

- Where are those spots predicted to contain L4 cells located in our subsetted data?

PART II: Joint analysis of two samples

1. Prepare the other slice

Repeat steps 2, 3 and 4 for the other sample. Name the tar-package as **mouse_posterior,** and the sample as **posterior1**. Open **QC_plots.pdf**.

-How does this sample look? Does it look very different?

2. Combine samples

Choose the two **seurat_obj_sctransform.Robj** objects and tool **Seurat v4 -Combine multiple samples**. As combining method, choose **Merge**. Run the tool. While the tool is running, study the manual of this tool.

- What are the two different combining methods, and how do they differ?
- Why do you think we used "merge" here?

3. PCA, clustering, and visualization for combined samples

Choose the **seurat_obj_multiple.Robj** from the previous step and run the tool **Seurat v4 - PCA, clustering, and visualization** for this combined object. Open **UMAP_plot.pdf**.
- How many clusters are there in this merged object?
- Are the clusters present in both samples?

4. Visualise gene expression of two samples

Choose **seurat_spatial_obj_pca.Robj** and tool **Seurat v4 -Visualise gene expression**. Type **gene name(s) = Hpca, Plp1** and run the tool. Open the resulting **Feature_plot.pdf**.
-Does it seem that the expression of these genes follows a pattern between the two images?

5. BONUS exercise: Repeat the analysis with "integrate" instead of "merge"

Repeat steps 2-4, but now in step 2, choose "integrate" as the combining method instead. Compare the clusters from merged and integrated analysis. Can you see any differences? When comparing the results you can use the Allen brain atlas as a reference for brain regions.

6. Send a support request to the Chipster team

In the top panel, click **Contact**. Click the **Contact support** button. In the new window,
-Click **Attach a copy of your last session** XXX
-Enter your **email address** and write a small **message.**

7. Share a session with a colleague (in this case with Eija and Maria)

Make sure that no file is selected. Go to the **Session info** panel, click the **three dots** next to the session name, and select **Share**. In the new window that opens
-click **Add rule**.
-In the **UserID** field, enter **jaas/support_session_owner**
-set **Rights** = **Read-only** (you don't want us to mess up your session)
-Click **Save** and **Close**.

Check what is your own UserID: Click on your **username** (top right corner) and select **Account**.

8. At home exercise: Repeat analysis with other samples

Download some other data from here and repeat the analysis for that sample.