Reproducibility in the PDAC-A

This file contains instructions for reproducing the results and figures contained in Real data analysis. We assume the working directory has been appropriately set.

**Step 1: "Step1\_Preprocessing.R"**

Preprocess raw data. The outputs of this step are "Real\_A\_author\_st1.RData"/ "PDAC\_A\_counts\_1.csv"/"PDAC\_A\_info\_1.csv".

In the "Real\_A\_author\_st1.RData ":

* spot.coor: The coordinates matrix of bulk ST data (each row represents a spot, the first column corresponds to its row index, the second column corresponds to its column index)
* Y: Preprocessed bulk ST data (each row represents a gene, and each column corresponds to a spot)
* W: Cell-type proportion matrix (each row stands for a spot and each column represents a cell type.
* K: Cell-type number

"PDAC\_A\_counts\_1.csv" and "PDAC\_A\_info\_1.csv" are inputs of SpatialDE and SOMDE, wherte the "PDAC\_A\_counts\_1.csv" stores spatial expression data, the "PDAC\_A\_info\_1.csv" contains the coordinates of spots.

**Step 2: "Step2\_CTSV.R****"**

Apply CTSV, SPARK-X, SPARK, BOOST-GP, trendsceek, SpatialDE, and SOMDE to the preprocessed data in Step1. The results are saved as ".csv" files.

**Step 3: "Step3\_Figure3A.R"**

Draw figure 1 and figure 5 in the manuscript using the output of step 2.

The **"SE\_gene\_A-1\_5.RData"** stores the SV genes of all the methods with an FDR rate of less than 0.01.