Overview of Pseudotime Analysis Application

Contents

1. Single Cell Analysis

Single cell analysis performed according to Seurat integration vignette (https://satijalab.org/seurat/articles/integration_introduction.html). FindClusters, FindConservedMarkers, FindMarkers and Differential Expression data for ggplot are repeated for each resolution. PHATE reduction is added to each resolution.

2. Psuedotime Analysis

2.1 Slingshot

For each resolution, the Seurat object is converted to single cell experiment and the dimensionality reduced. Slingshot is then run using the Seurat cluster labels. It is run for both UMAP and PHATE. The outputs from Slingshot are converted to slingshot datasets.

An object is also produced for each resolution, for each possible starting cluster, for both UMAP and PHATE.

heatmaps

For each resolution, the Seurat package is used to discern Variable Features. The p/q-value for these variable genes are calculated using gam from the package gam. The results are sorted by q-value. The heatmaps are annotated with clusters and lineages. Heatmaps are produced for UMAP and PHATE

2.2 tradeSeq

For each resolution, the clusters and counts are extracted from the single cell Slingshot output and Seurat object, respectively. The counts are filtered to remove those with a count >1 in >=120 cells. This significantly speeds up the code and was recommended by the developer. Next, fitGAM is used to produce the final output of tradeSeq with the knots parameter set to 5.

Heatmaps

For each resolution, associationTest is used to assess whether gene expression is associated with pseudotime. They are sorted by P-value (due issues with q-value in the tradeSeq package). The heatmaps are annotated with clusters and lineages. Heatmaps are produced for UMAP and PHATE

2.3 Monocle 2

For each resolution, count, phenotype and feature data are extracted from the Seurat object to produce a cell data set. Next, standard Monocle 2 pipeline is carried out (estimateSizeFactors, estimateDispersions, dispersionTable, setOrderingFilter) to prepare the data for pseudotime analysis. The dimension is reduced using DDRTree and the cells are ordered.

Heatmaps

For each resolution, genes are tested for those that have differential expression as a result of pseudotime. They are then sorted by q value and the top 40 are selected. Clusters are annotated on the heatmap.

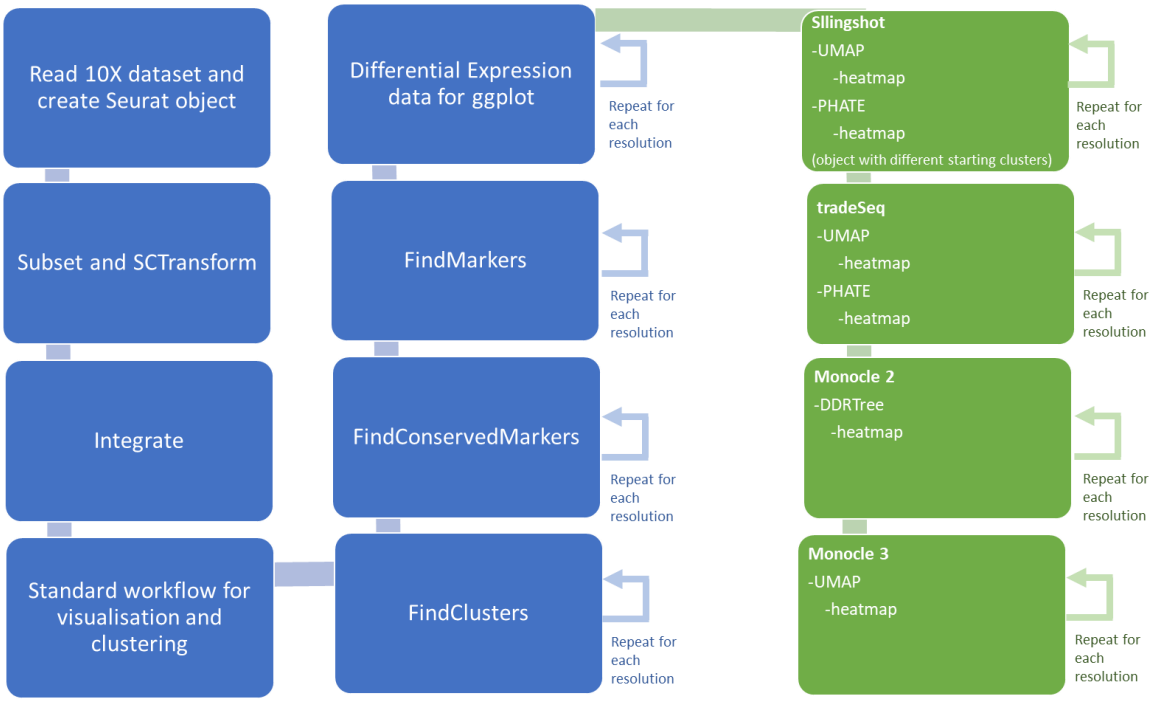
2.4 Monocle 3

For each resolution, the Seurat object is converted to a cell data set. Cells are then clusters and subset. learn_graph learns the principal graph. Cells are finally ordered.

Heatmaps

For each resolution, graph_tests for genes differentially expressed as a function of pseudotime. They are ordered by q-value and the top 40 selected. lusters are annotated on the heatmap.

3. Overview Figure of Workflow



4. Overview Figure of Tabs