

Entropy (might do this later)

<https://rdrr.io/github/aet21/SCENT/f/inst/doc/SCENT.Rmd>

PPI:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2048726/bin/plntphys_pp.107.103465v2_index.html

SCENIC:

http://htmlpreview.github.io/?https://github.com/aertslab/SCENIC/blob/master/inst/doc/SCENIC_Setup.html

How to Calculate Correlation Between Continuous & Categorical Variables

I.e. correlation between counts/number of features and ploidy/...

Point biserial correlation: <https://www.statology.org/correlation-between-continuous-categorical-variables/>

Subclustering

Might be beneficial for clusters with high uncertainty and to avoid cluster names with lots of different cell types

Unsupervised: <https://satijalab.org/seurat/reference/findsubcluster>

Supervised: e.g. based on specific markers (as they did for distinguishing QC and columella in Denyer paper, fig. 4)

GO-enrichment of clusters

<https://cran.r-project.org/web/packages/gprofiler2/vignettes/gprofiler2.html>

<https://bioconductor.org/packages/release/bioc/html/goseq.html>

Cell-type specific networks:

<https://github.com/VIB-PSB/MINI-EX>

Single-cell data simulation (could also skip this and start on small cluster)

<https://github.com/dynverse/dyngen>

https://dyngen.dynverse.org/articles/getting_started.html

https://dyngen.dynverse.org/articles/showcase_backbones.html

Trajectory inference

<https://github.com/dynverse/dynguidelines>

dynguidelines Tutorial Citation

Benchmark study Evaluating methods with dynbenchmark Part of

▼ Topology ADAPTED

Do you expect multiple disconnected trajectories in the data?
☐ Yes ☐ I don't know ☐ No

Do you expect a particular topology in the data?
☐ Yes ☐ No

Do you expect cycles in the data?
☐ Yes ☐ I don't know ☐ No

Do you expect a complex tree in the data?
☐ Yes ☐ I don't know ☐ No

▼ Scalability ADAPTED

Number of cells

Number of features (genes)

Show code </> Show/hide columns Options Close & use

Lenses	Default	Summary (Fig. 2)	Method	Scalability	Stability	Usability	Accuracy	Overall	Everything
Method				Accuracy		Scalability			
</>	Name	Priors	Errors	Tree	Overall				
✓ </>	PAGA Tree	×		100	99	38s	2GB		
✓ </>	PAGA	×		99	89	35s	2GB		
✓ </>	MST			88	89	13m	3GB		
✓ </>	Slingshot			82	100	23m	6GB		
</>	pCreode			79	78	36m	7GB		
</>	ElPiGraph			74	54	3h	10GB		
</>	MERLOT	×		66	16	1d	4GB	▲ Errors	
</>	SCUBA			65	70	2h	7GB		
</>	URD	×		61	43	60m	5GB	▲ Errors	
</>	Mpath	×		56	60	5h	9GB		
</>	cellTree			52	69	21m	6GB		

<http://bioconductor.org/books/3.13/OSCA.advanced/trajectory-analysis.html>

<https://bioconductor.org/packages/devel/bioc/vignettes/slingshot/inst/doc/vignette.html>

Differential expression along pseudotime

<https://bioconductor.org/packages/release/bioc/vignettes/tradeSeq/inst/doc/tradeSeq.html>

<https://kstreet13.github.io/bioc2020trajectories/articles/workshopTrajectories.html>

<https://statomics.github.io/tradeSeq/articles/tradeSeq.html>

Probabilistic methods

A descriptive marker gene approach to single-cell pseudotime inference (Campbell et al., 2019):

<https://github.com/kieranrcampbell/ouija>

Order Under Uncertainty: Robust Differential Expression Analysis Using Probabilistic Models for Pseudotime Inference (Campbell et al., 2016): <https://github.com/kieranrcampbell/pseudogp>