Something that we could study a bit more in depth (for example as a 'research' aim) as the ploidy level of cells seems to match our putative trajectories is a phenomenon called endoreduplication (basically duplication of the nuclear genome):

https://en.wikipedia.org/wiki/Endoreduplication

Two background papers: https://doi.org/10.1016/j.tplants.2014.11.007, https://doi.org/10.1105/tpc.17.00983

We could use trajectory inference to detect ploidy associated gene expression profiles and gene regulatory networks (although this is a bit tricky since we use genetic markers for assigning ploidy of cells a priori)

Discarded cell types

Can we somehow plot discarded cells on UMAP?

"The biggest practical concern during QC is whether an entire cell type is inadvertently discarded. There is always some risk of this occurring as the QC metrics are never fully independent of biological state."

http://bioconductor.org/books/3.16/OSCA.advanced/quality-control-redux.html#qc-discard-cell-types

Cell-cycle assignment

Plot Shahan2020 cell-cycle markers on UMAP

http://bioconductor.org/books/3.16/OSCA.advanced/cell-cycle-assignment.html

Automated cell-type annotation

Dot plot with curated markers (e.g. Shahan, 2022) + UMAP, distribution of cell types per cluster (pericycle = phloem pole + xylem pole)

Besides 'manual' cluster annotation with known cell-type markers there are statistical methods that are automated either, conditional on the cluster annotation or cluster free (reference paper: https://doi.org/10.1016/j.csbj.2021.01.015):

Marker gene based:

https://github.com/ZJUFanLab/scCATCH

https://irrationone.github.io/cellassign/articles/introduction-to-cellassign.html

Correlation based: (especially for the developmental zones: meristem, elongation, maturation)

https://github.com/Hsu-Che-Wei/COPILOT/blob/master/jupyter_notebook/1-Correlation_Based_Annotation.ipynb

GO-enrichment of clusters

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

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

Single-cell data simulation

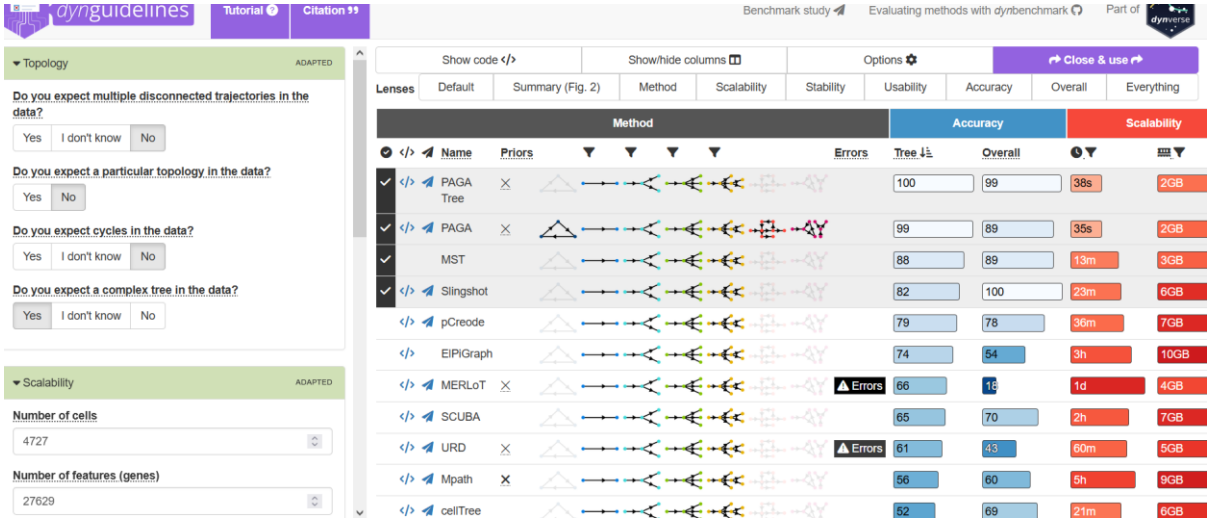
https://github.com/dynverse/dyngen

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

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

Trajectory inference

https://github.com/dynverse/dynguidelines



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

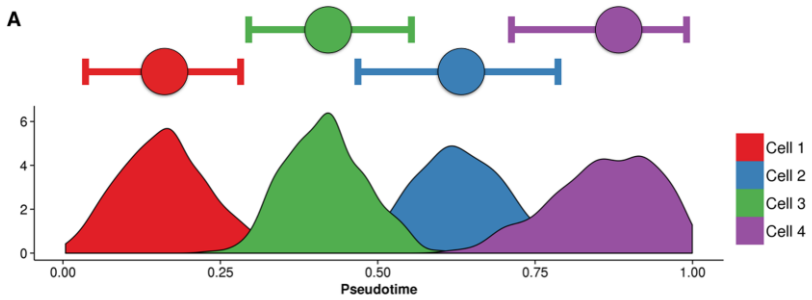
https://bioconductor.org/packages/release/bioc/vignettes/tradeSeq/inst/doc/tradeSeq.html

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

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


A
6
4
2
0
0.00
0.25
0.50
0.75
1.00
Pseudotime
Cell 1
Cell 2
Cell 3
Cell 4