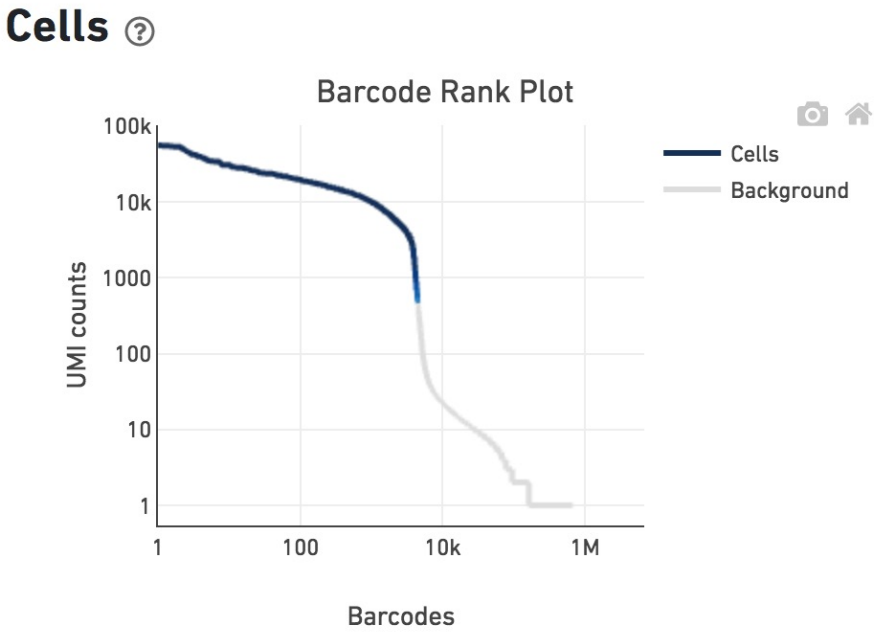
Single cell transcriptomics

Quality control

Cell calling

- Doublets: high #UMI/cell
- Dropouts: low #UMI/cell



Further cell filtering

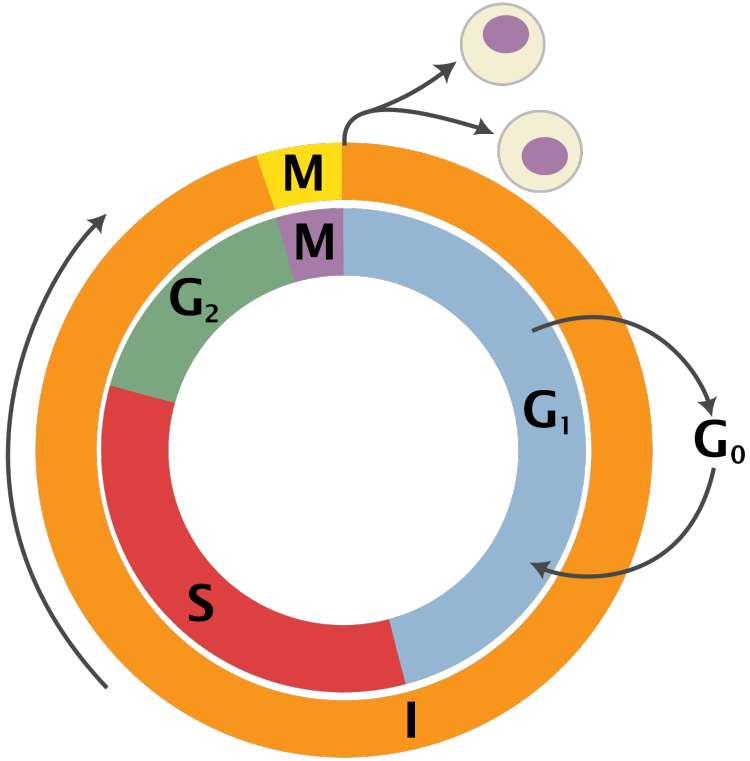
- % UMI in mitochondrial genes:
  - Points to dying/stressed cells
  - Filter by threshold (e.g. 20%)
- Number of detected genes
  - Can also point to dropouts/doublets

Confoundig variables

- % mitochondria
- % ribosomal genes
- % dissociation genes
- Library size
- Sex
- Age
- Cell cycle
- ..

Cell cycle

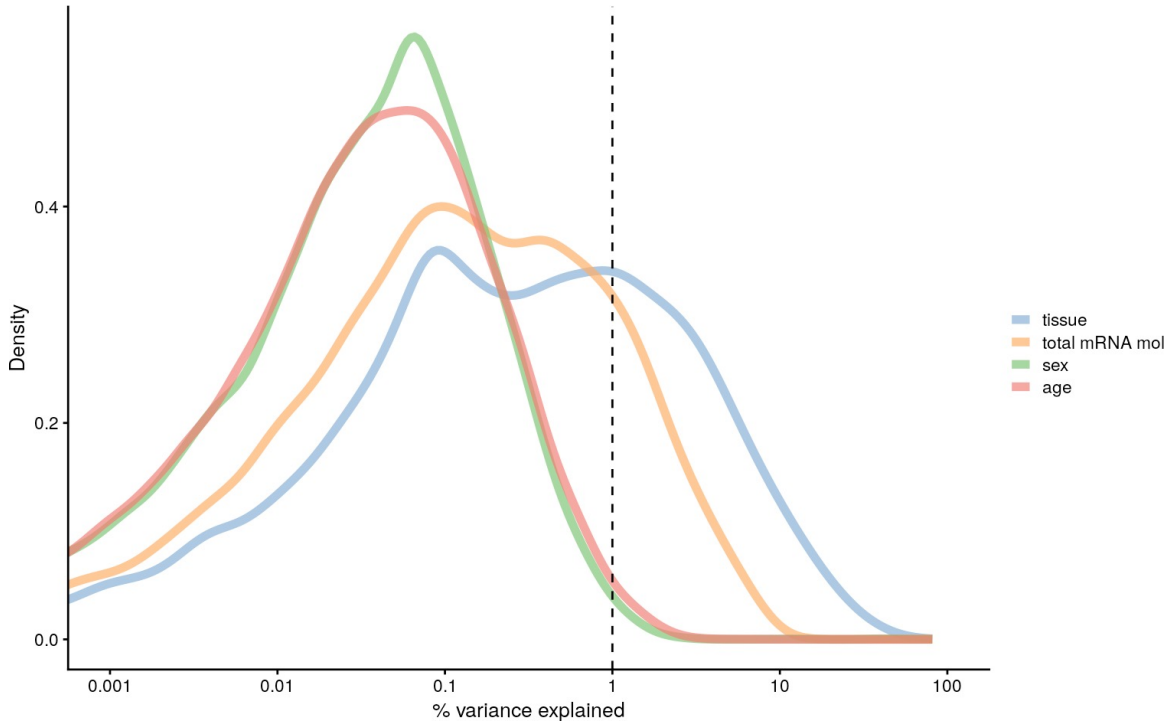
- Default clustering might be according to state
- Cycling state can be regressed out for clustering according to cell type



Cell_Cycle_2.png: Original uploader was Zephyris at en.wikipedia
https://commons.wikimedia.org/w/index.php?curid=12800954

Find explanatory variables

- With the package `scater`
- Plot variance explained per gene



Regress out for clustering

- Regress out variables explaining a lot of variation
- At the scaling step -> affects PCA
- **OR**: integrate data (day 2)