Interactive scRNA-Seq analysis with the Single Cell Toolkit (SCTK)

W. Evan Johnson, Ph.D.
Boston University
@wejlab
ASSIGN: Pathway activity profiling

Single Cell Toolkit: interactive single cell analysis
PathoScope: microbiome and pathogen identification
ComBat and BatchQC: batch correction
Single Cell RNA-seq
(The Future of RNA-sequencing)
Applications for Single Cell Sequencing:

- **Cancer**: intra-tumoral heterogeneity
- **Development**: characterize every cell in blastocyst
- **Other applications**: novel/rare cell type discovery
- **Infectious Diseases**: combine with metagenomics to explore host response (e.g. TB)
- And more!

The Future of RNA-sequencing
Tumor Heterogeneity

Metastatic breast tumor
Pre-treatment

Metastatic breast tumor
Post-treatment

Metastatic breast tumor
Refractory cancer

Course of treatment
Pathway Activity in Breast Cancer

Critical Gaps in scRNA-seq data

• Complex data
  • New analysis challenges

• Interactive, Simple Analysis?
  • Inexperienced Users
  • Optimizing Parameters
  • Filtering Failed Samples
  • Filtering Low expression genes

• Some packages already exist
  • QC
  • Clustering
  • Full Analysis Portal?
SingleCellITK
Single Cell Toolkit

• Standard R package *with* a Shiny toolkit

• R functions on top of a SingleCellExperiment (SCE) object
  • Can download analysis performed in the toolkit and continue on the command line

• SCE object can be brought in/out of Shiny at any stage
  • Great for common tasks:
    • Interactive clustering/visualization
    • Differential gene/pathway analysis

• (Also works for bulk RNA-seq analysis)
SummarizedExperiment

```
se <- SummarizedExperiment(
  assays,
  rowData,
  colData,
  exptData
)
```
SingleCellExperiment

- `reducedDims`
  - PCA, TSNE, any other sample size matrix of dimensionality reduction data
- `isSpike` – indicate which probes are spike ins
- `sizeFactors` – add scaling factors
- Named assays: counts, normcounts, logcounts, cpm, tpm
  - Just a convention
- `DelayedArray` – run common array functions without loading it into memory
Shiny

- Simple Web Apps Written in R
- Interactive and Reactive
- Customizable, Modular

Iris k-means clustering

![Iris k-means clustering graph](image-url)
singleCellTK

Interactive Analysis of Single Cell RNA-Seq Data

Bioconductor version: Release (3.9)

Run common single cell analysis directly through your browser including differential expression, downsampling analysis, and clustering.

Author: David Jenkins

Maintainer: David Jenkins <dfj at bu.edu>

Citation (from within R, enter citation("singleCellTK")):

Single Cell Toolkit
Filter, cluster, and analyze single cell RNA-Seq data
Need help? Read the docs.
Single Cell Toolkit
Filter, cluster, and analyze single cell RNA-Seq data
Need help? Read the docs.

Upload
(help)

Successfully Uploaded!

Choose data source:
- Upload files
- Upload SCtkExperiment RDS File
- Use example data

Choose Example Dataset:
mouseBrainSubset
Single Cell Toolkit

Filter, cluster, and analyze single cell RNA-Seq data

Need help? Read the docs.

Upload

(帮)

✔ Successfully Uploaded!

Choose data source:
- Upload files
- Upload SCiteExperiment RDS File
- Use example data

Choose Example Dataset:
mouseBrainSubset
SCTK features – **Differential Expression analysis - MAST**

**MAST**

Select Assay:
- logcounts

**Adaptive Thresholding:**
- Run Thresholding

**Hurdle Model:**
- Use Adaptive Thresholds

Select fold change threshold
- 0
- 0.5
- 1

Select expression threshold
- 0
- 1

Select Condition for Hurdle Model
- condition

p-value (FDR) cutoff:
- 0.01
- 0.05
- 0.2

Run DE: Use Hurdle
SCTK features – Pathway activity analysis - GSVA
Thanks

• Johnson Lab
  • David Jenkins
  • Mohammad Khan
  • Tyler Faits
  • Yuqing Zhang
  • Yue Zhao

• Additional Toolkit Contributors
  • Emma Briars
  • Sebastian Carrasco Pro
  • Steve Cunningham
  • Sean Corbett

• Bild Lab (University of Utah)
  • Andrea Bild
  • Mumtahena Rahman
  • Shelley Macneil
  • Sam Brady

• Single Cell Workgroup
  • Josh Campbell
  • Masanao Yajima
Common Questions

• How is this different from other tools/packages?
  • (Is this the right question?)

• Can we import data to/from other tools?
  • CellRanger object
  • To/from Seurat

• Is it scalable to larger datasets?