Single cell RNAseq analysis in R

22-23 August 2022

Resources shared in Slack

Resource	Link
Website for checking that your gene id is correct. For example, "c-myc" is my favourite gene but the HGNC id for this gene is "MYC"	https://www.genecards.org/
VDJ analysis: Note: neither of these ecosystems/packages use Seurat,but should work with 10x VDJ output	Immcantation https://immcantation.readthedocs.io/en/stable/ scRepertoire https://www.bioconductor.org/packages/release/bioc/vignettes/scRepertoire/inst/doc/vignette.html
Bioconductor has a package to facilitate automatic outlier identification via PCA of nFeature, nCount, mt.percent. But the downside is that it's got a stochastic nature and the outliers may differ.	http://bioconductor.org/books/3.15/OSCA.basic/quality-control.html#identifying-low-quality-cells
Doublets Here is one of tools that works on Seurat object.	https://www.cell.com/cell-systems/pdf/S2405-4712(19)30073-0.pdf https://github.com/chris-mcginnis-ucsf/DoubletFinder
Resources for spatial transcriptomics. Note this is quite different to scRNA-seq	https://lmweber.org/OSTA-book/
Understanding UMAPs	https://pair-code.github.io/understanding-umap/

Making sense of PCAs	https://stats.stackexchange.com/questions/2691/making-sense-of-principal-component-analysis-eigenvectors-eigenvalues
Seurat vignette on removing cell cycle genes	https://satijalab.org/seurat/articles/cell_cycle_vignette.html
This is a method for comparing resolutions.	https://cran.r-project.org/web/packages/clustree/vignettes/clustree.html
code to look at several resolutions at once `pbmc <- FindClusters(pbmc, resolution = c(0.5, 0.1,0.2,0.3,0.4,0.8,1,1.2,5)) DimPlot(pbmc, reduction = 'umap', group.by =c('RNA_snn_res.0.5','RNA_snn_res.0.1', 'RNA_snn_res.1' ,'RNA_snn_res.5'))`	
How to check the validity of clusters	https://www.biorxiv.org/content/10.1101/2022.08.01.502383v1
Significance analysis for clustering single cell RNAseq data	https://github.com/igrabski/sc-SHC
On the Seurat default clustering algorithm (Louvain) and an alternative (Leiden)	https://www.nature.com/articles/s41598-019-41695-z
Comparison of various different tests for marker genes	https://www.biorxiv.org/content/10.1101/2022.05.09.490241v1
Resources for finding reference sets and public scRNAseq data	GEO: https://www.ncbi.nlm.nih.gov/geo/ Single Cell Expression Atlas https://www.ebi.ac.uk/gxa/sc/home Similar to EMBL's repository, I've found this spreadsheet quite helpful in my work: www.nxn.se/single-cell-studies/gui Associated with this paper: https://doi.org/10.1093/database/baaa073

	Broad Institute https://singlecell.broadinstitute.org/single_cell Seurat has some datasets too https://github.com/satijalab/seurat-data CELLxGENE has quite a collection https://cellxgene.cziscience.com/collections
CellSNP-lite to for the snps calling and genotyping:	https://github.com/single-cell-genetics/cellsnp-lite