An Introduction to Single-Cell Genomics

(Actually, mostly transcriptomics)

(Actually, mostly from 10X Genomics)

Bastien JOB

bastien.job@gustaveroussy.fr

INSERM / Gustave Roussy

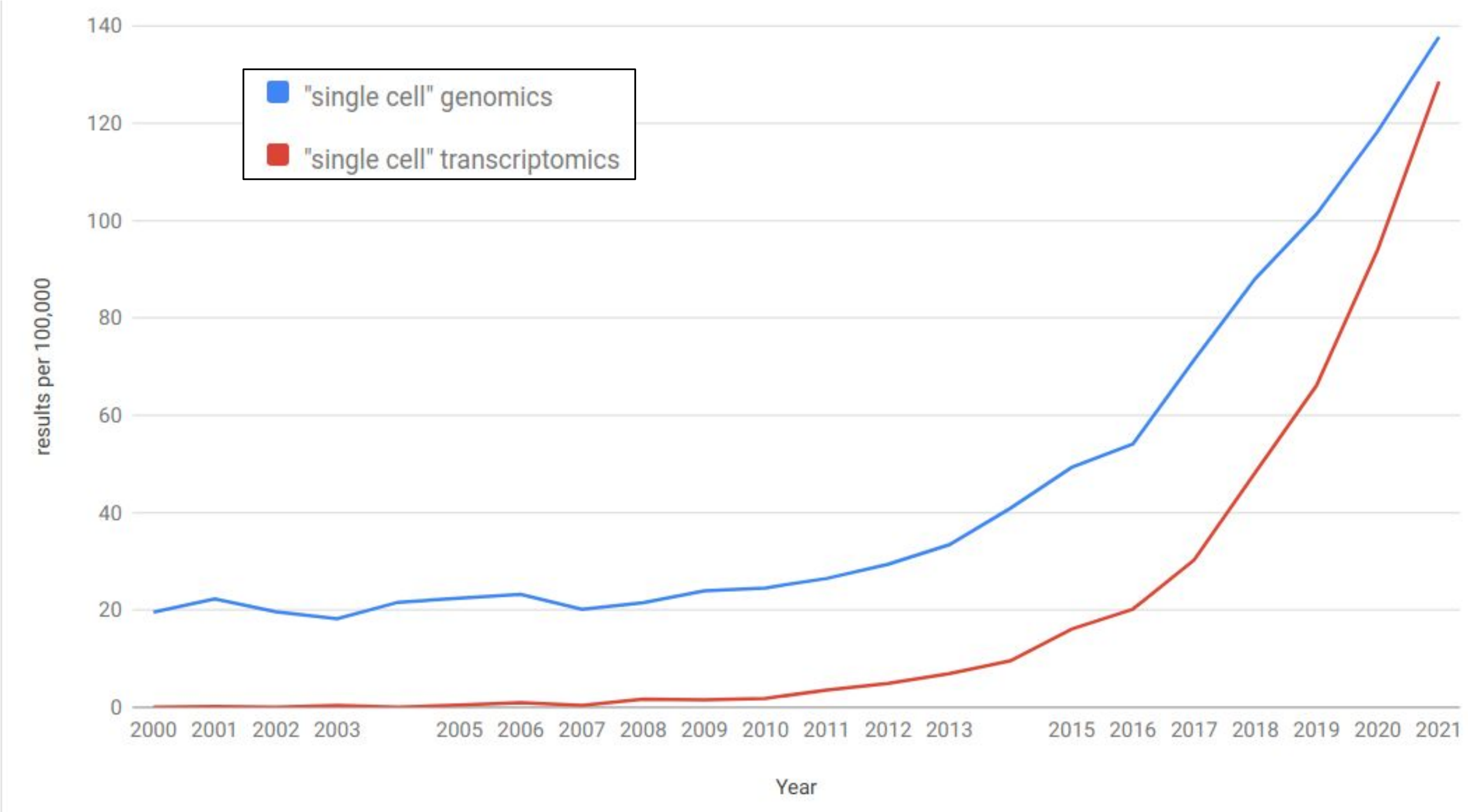
IFSBM 2021/11/22

So you say you've heard

About

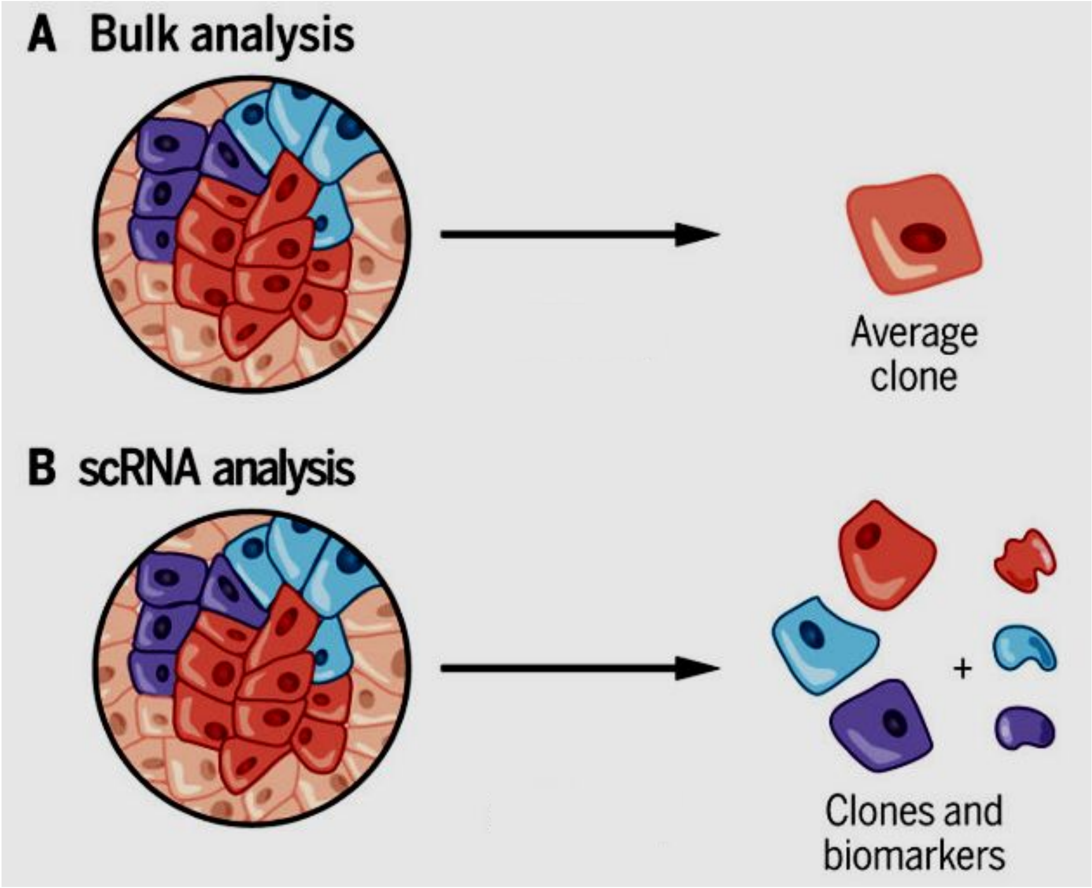
Single cell ?

Single cell in peer-reviewed publications (2021)



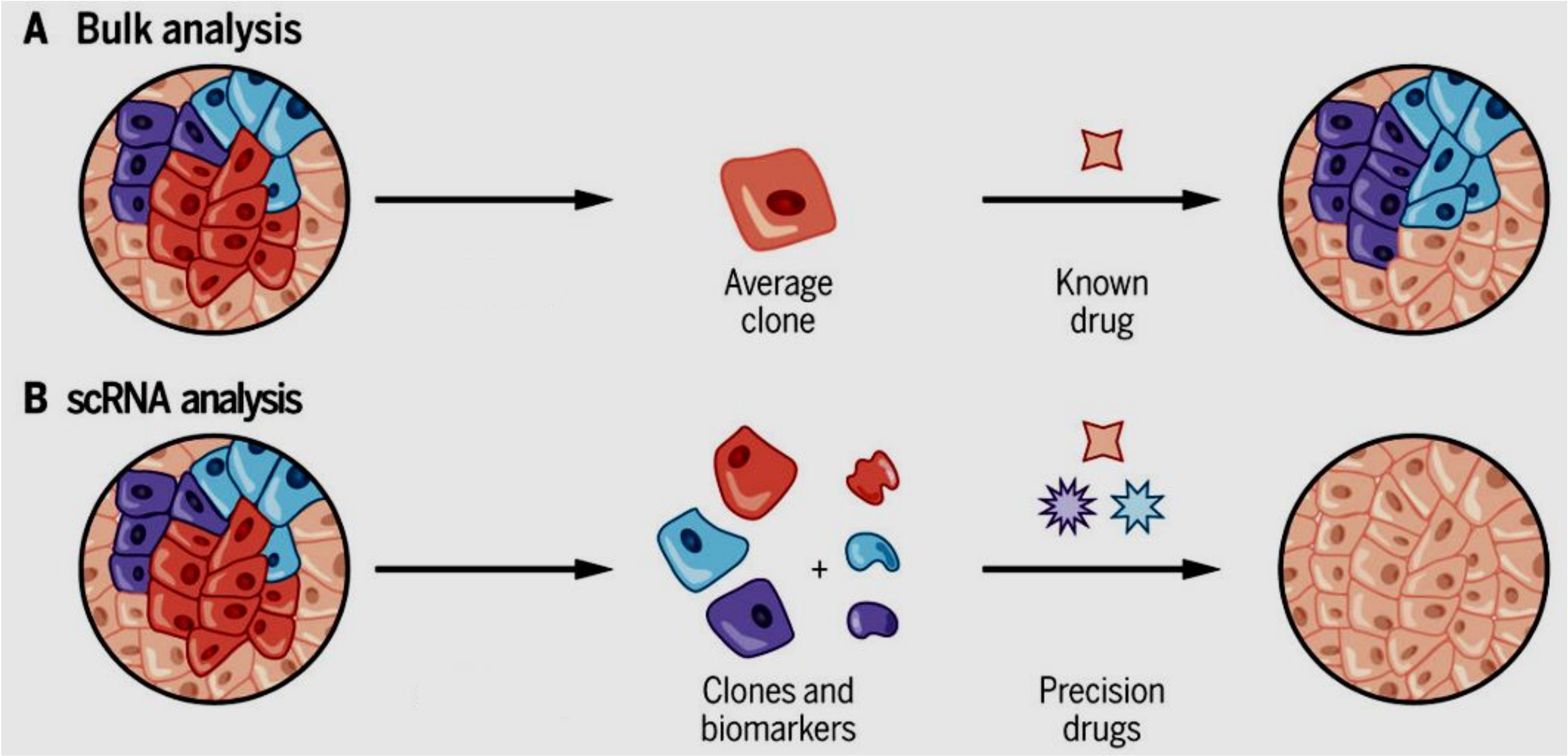
https://esperr.github.io/pubmed-by-year/

Why so much hype ?



Shalek and Benson, 2017

Why so much hype ? (pathology)



Shalek and Benson, 2017

Why so much hype ?

Bulk

Single cell

Spatial single cell

Based upon Shalek and Regev, 2016

Why so much hype ?

Bulk

Single cell

Spatial single cell

Full-length spatial single cell

Based upon Shalek and Regev, 2016

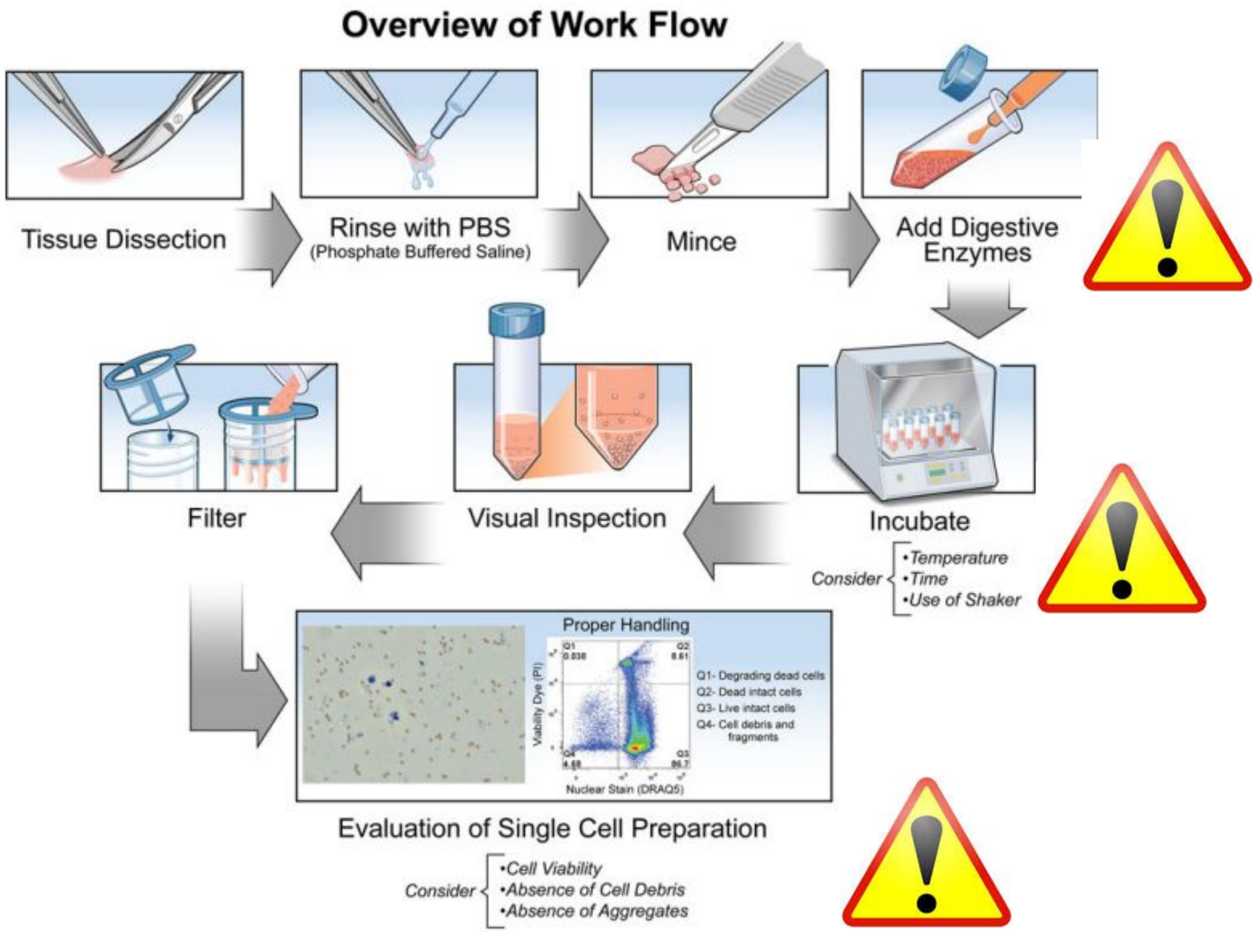
From

Broad tissue

To

Isolated cells

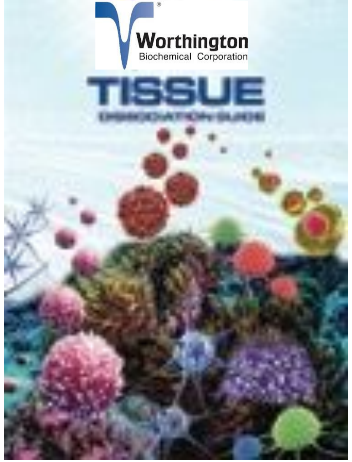
Cells health and dissociation



Reichard & Asosingh, Cytometry (2018)

Cells health and dissociation : *Worthington* helpdesk

1. Type of tissue
2. Species of origin
3. Age of the animal
4. Genetic modification(s) (knockouts, etc.)
5. Dissociation medium used
6. Enzyme(s) used
7. Impurities in any crude enzyme preparation used
8. Concentration(s) of enzyme(s) used
9. Temperature
10. Incubation times



Tissue Tables (references, grouped by tissue type and species)

Adipose/Fat	Adrenal	Bone	Brain
Cartilage	Colon	Endothelial	Epithelial
Eye	Heart	Intestine	Kidney
Liver	Lung	Lymph nodes	Mammary
Miscellaneous	Muscle	Neural	Pancreas
Parotid	Pituitary	Prostate	Reproductive
Scales	Skin	Spleen	Stem
Thymus	Thyroid/Parathyroid	Tonsil	Tumor

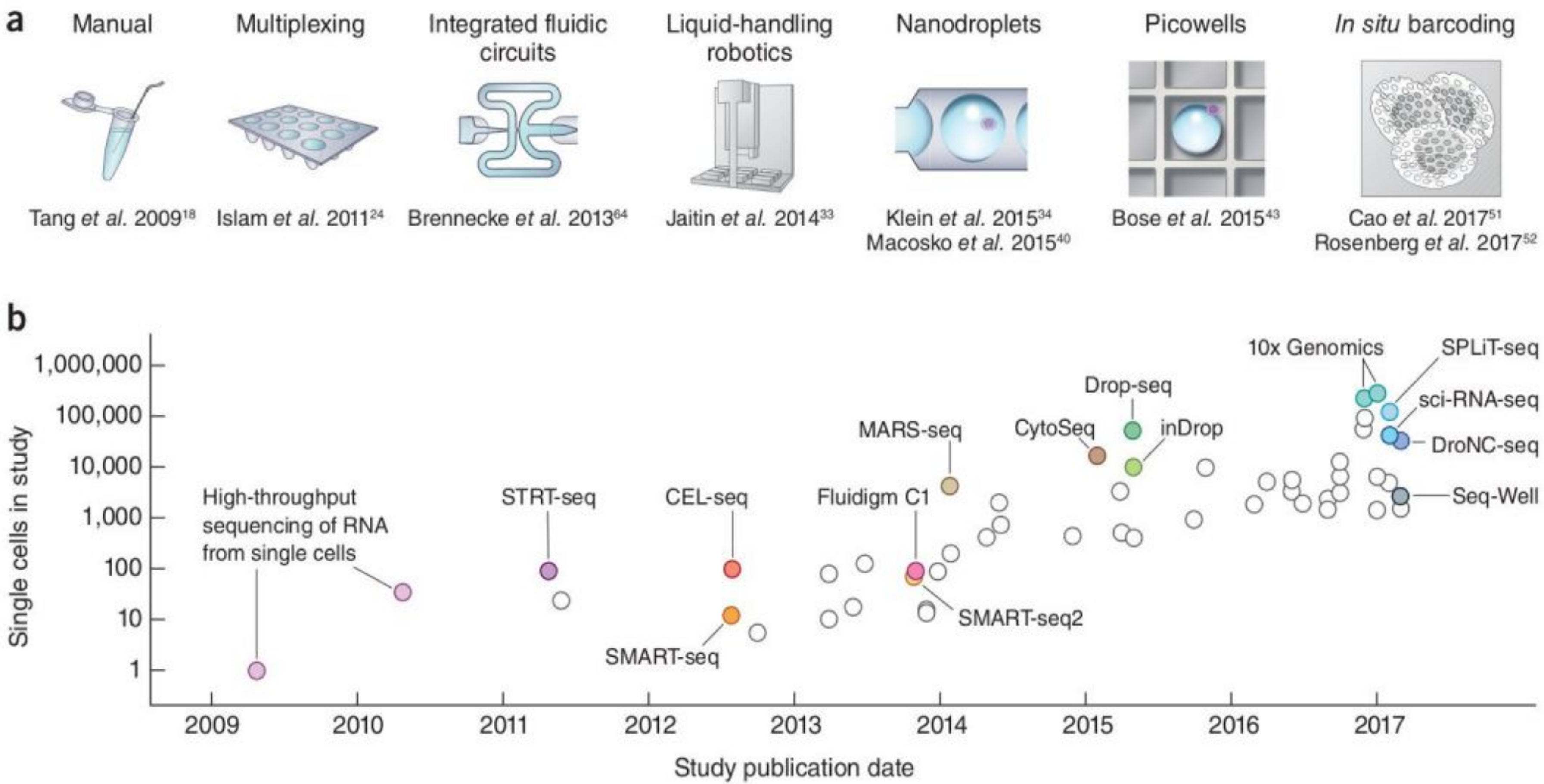
II. Cell Isolation Theory
- Tissue Types
  - Epithelial Tissue
  - Connective Tissue
- Dissociating Enzymes
  - Collagenase
  - Trypsin
  - Elastase
  - Hyaluronidase
  - Papain
  - Chymotrypsin
  - Deoxyribonuclease I
  - Neutral Protease (Dispase)
  - Trypsin Inhibitor
  - Animal Origin Free (AOF) Enzymes
  - Celase® GMP

III. Cell Isolation Techniques
- Methods & Materials
  - Working With Enzymes
  - Basic Primary Cell Isolation
  - Equilibration with 95%O_2:5%CO_2
  - Trituration
  - Enzymatic Cell Harvesting
  - Cell Adhesion and Harvesting
  - Trypsin for Cell Harvesting
  - Cell Release Procedure
- Optimization Techniques
  - General Guidelines
  - Optimization Strategy
  - Cell Quantitation
  - Measure of Viability

IV. Use-Tested Cell Isolation Systems

http://www.worthington-biochem.com/tissuedissociation/default.html

Cells isolation : technologies over the last decade



Svensson et al. Nature protocols (2018)

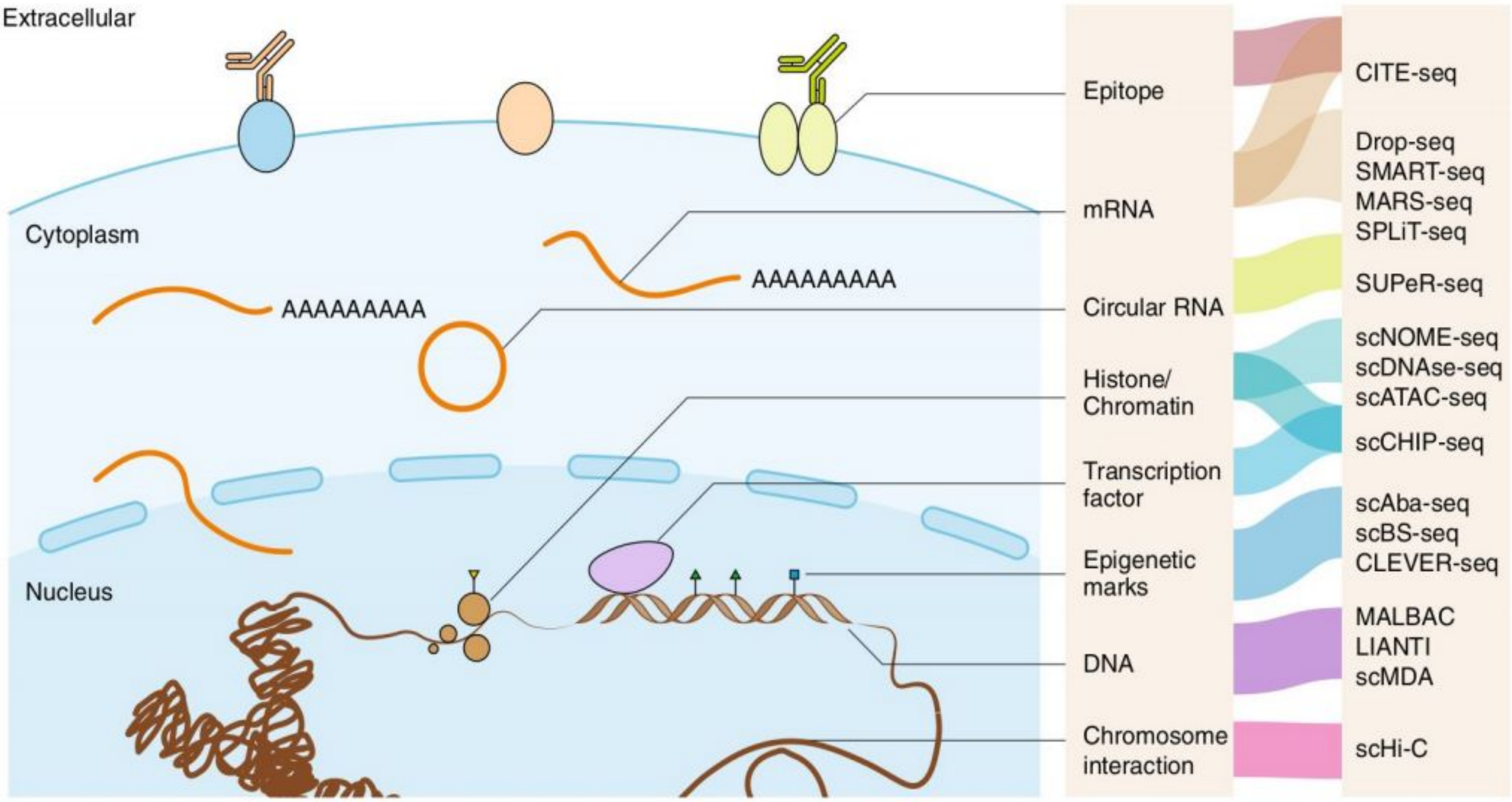
From

Isolated cells

To

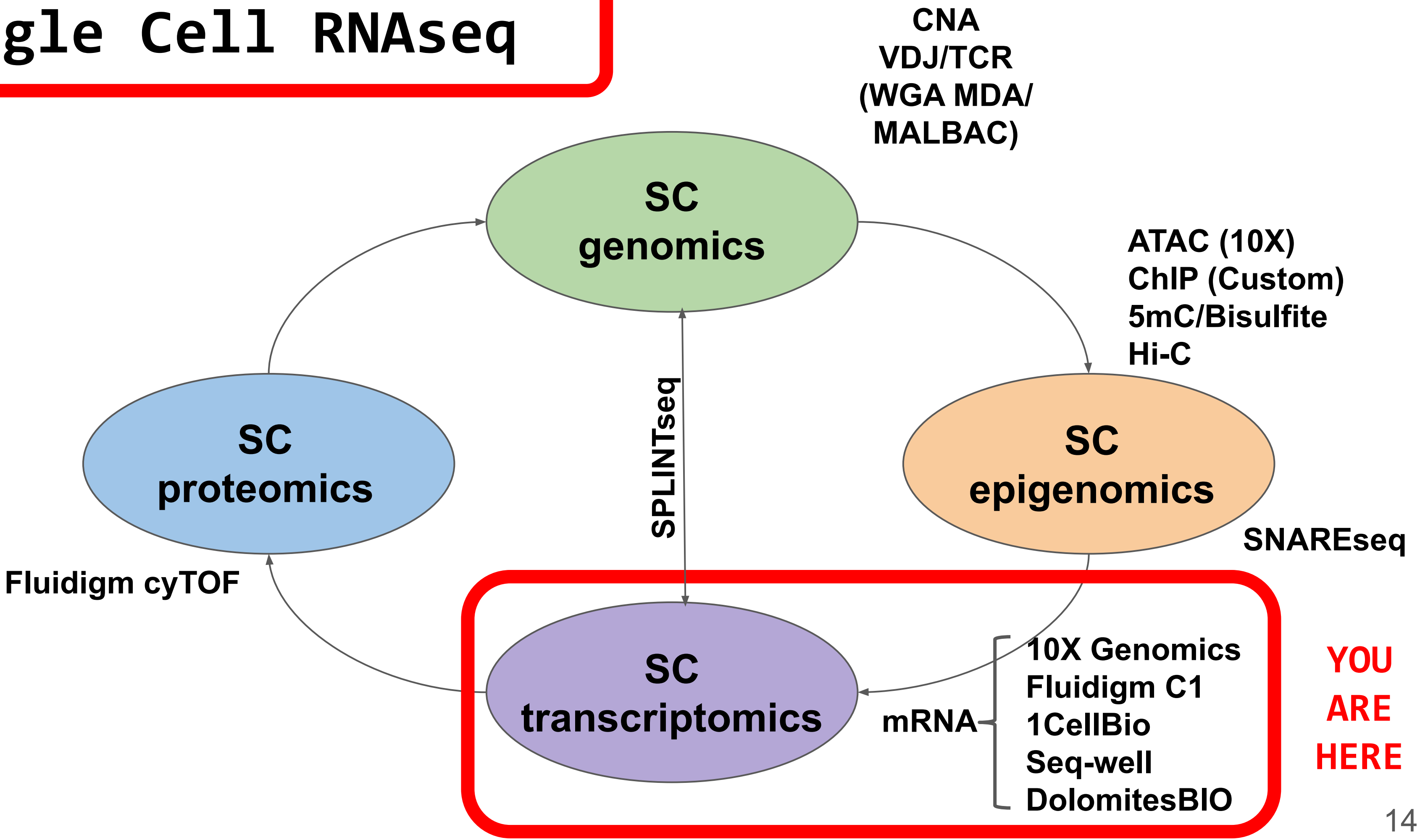
Nucleotide sequences (reads)

Several protocols for several purposes

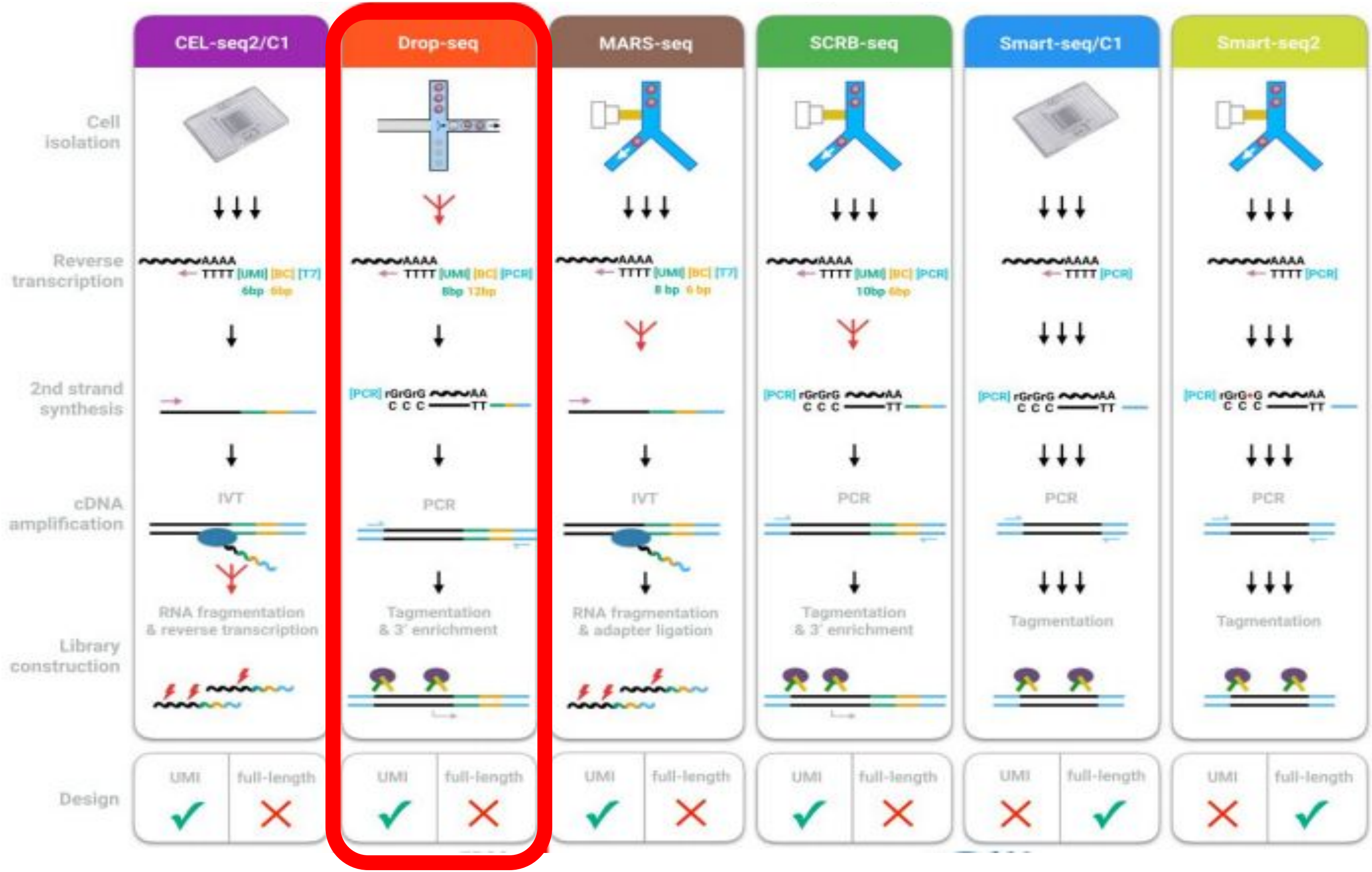


Ren et al. Genome Biology (2018)

Single Cell RNAseq



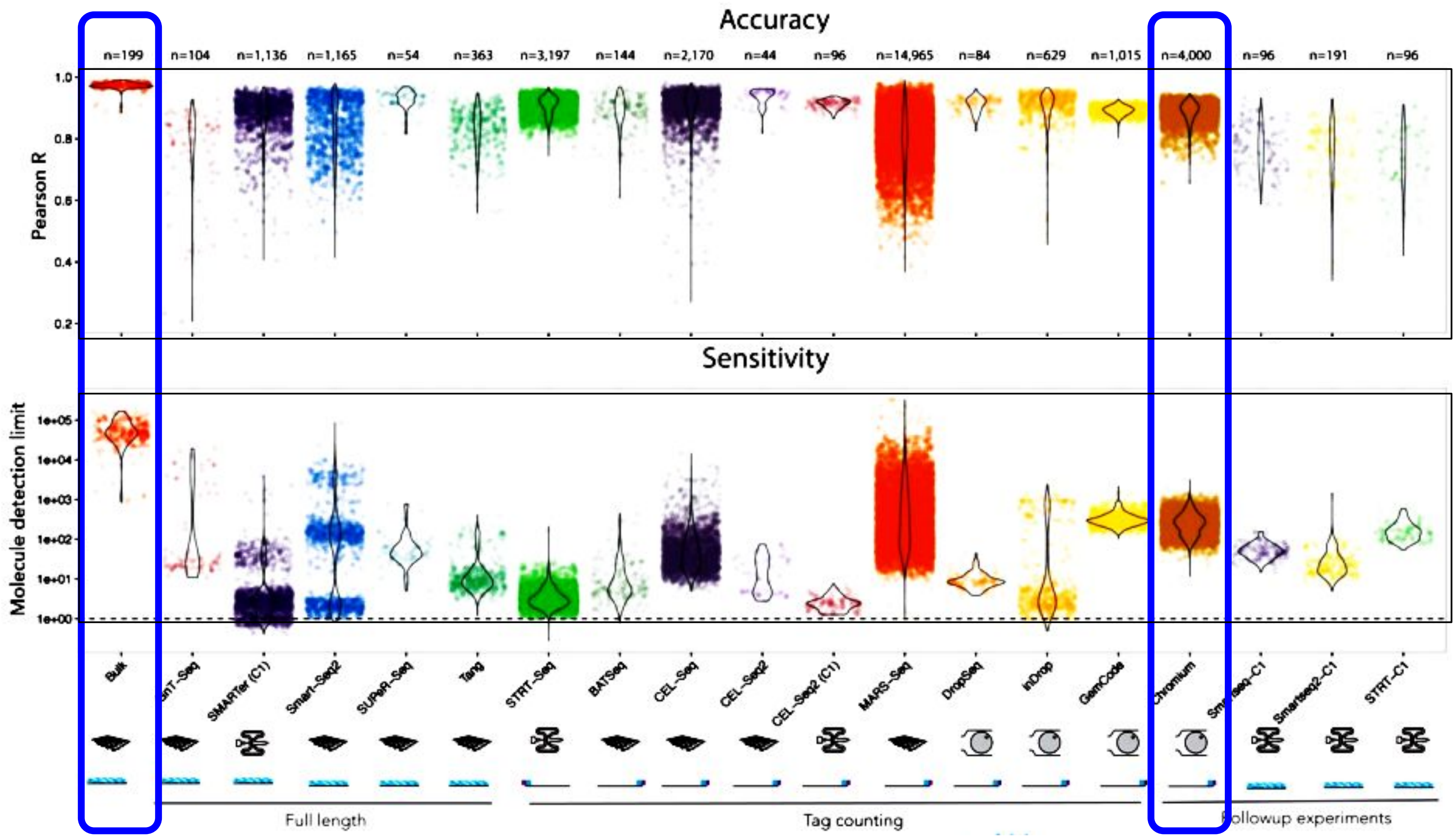
From isolated cells to sequences (Drop-seq / 10X)



Ziegenheim et al. Molecular Cell (2017)

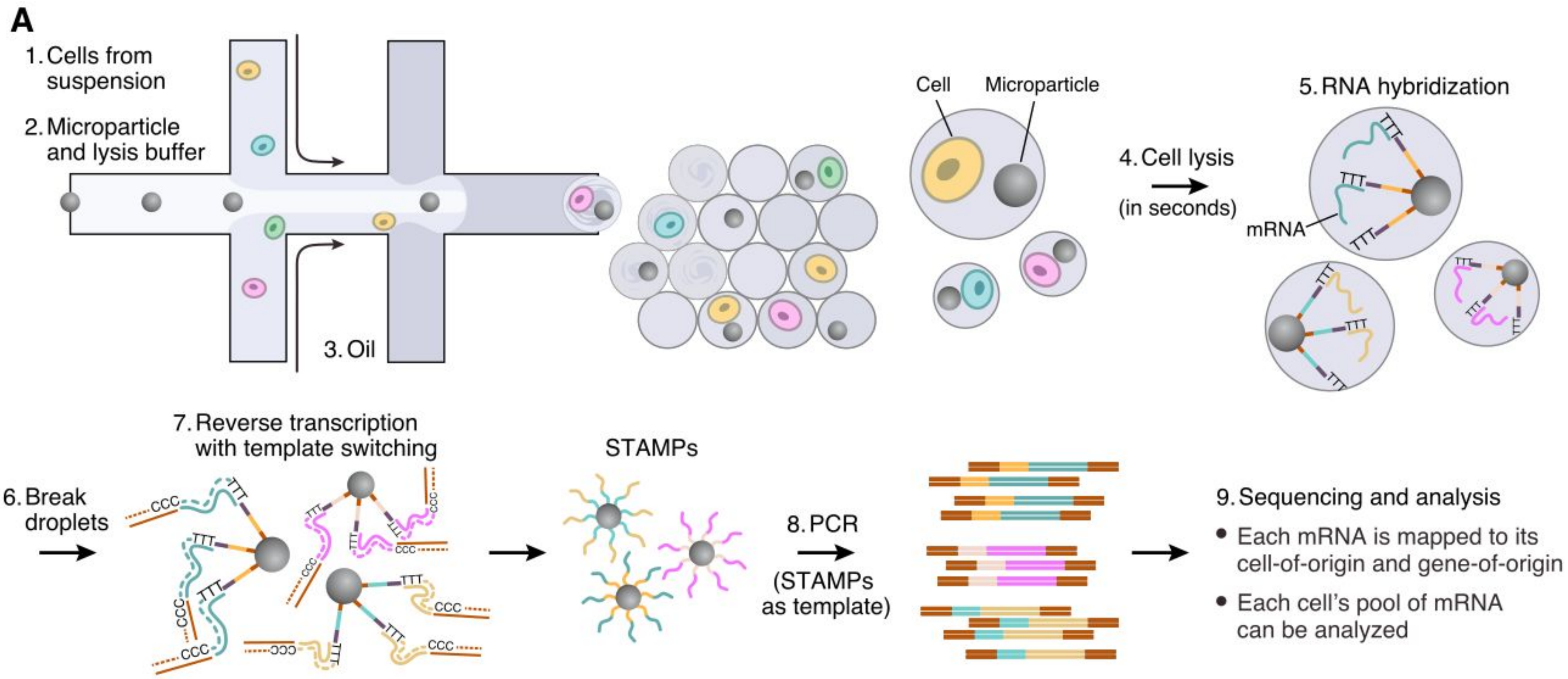
Bulk

10X



Svensson et al. Nature methods (2017)

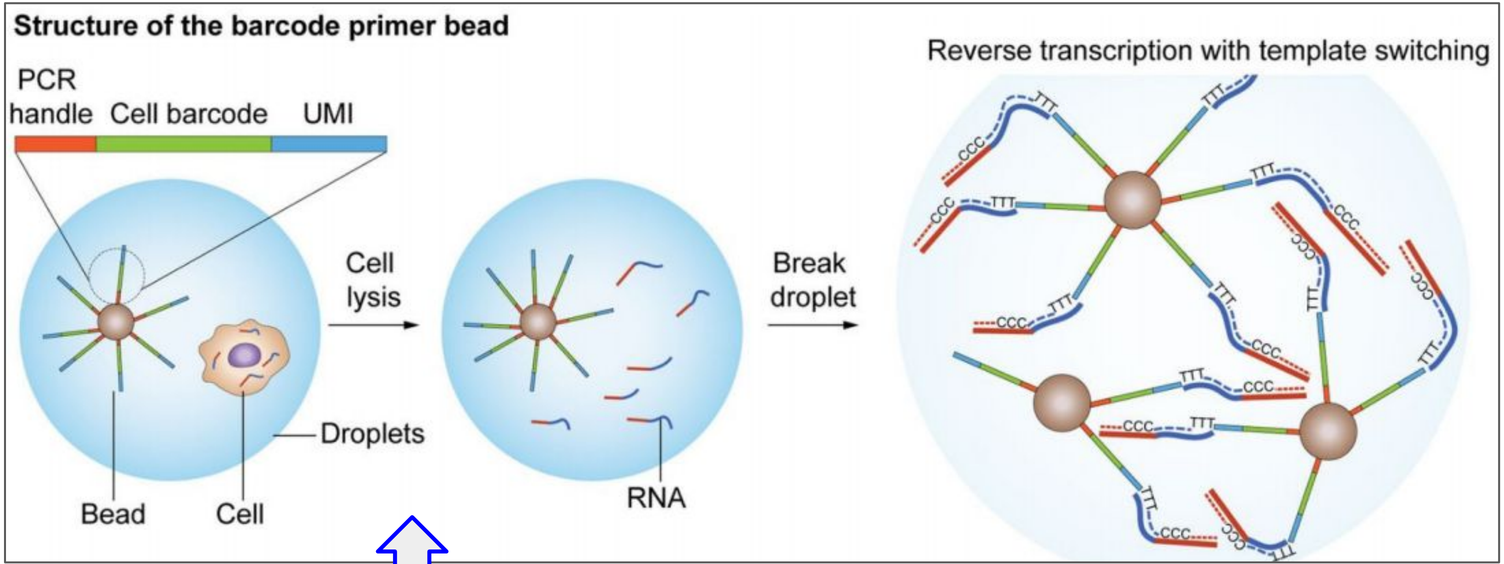
Drop-seq



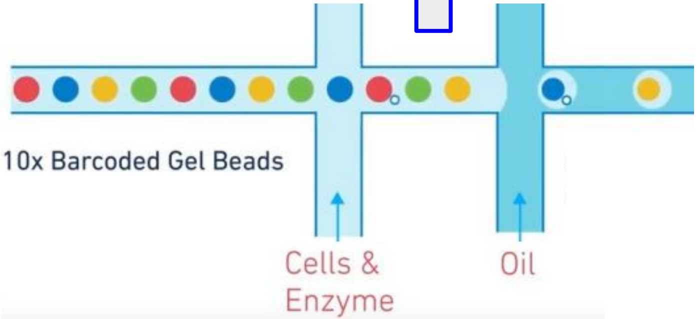
Macosko 2015

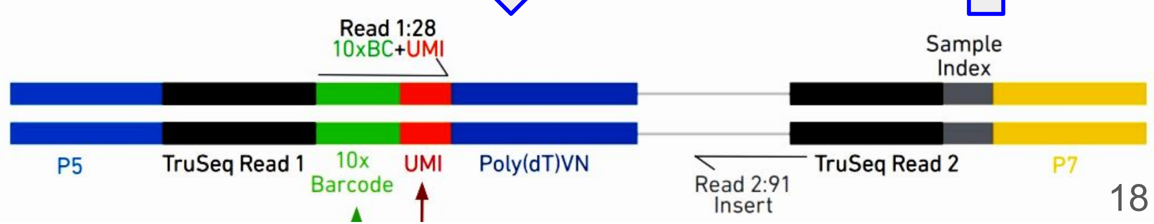
STAMPs = single-cell transcriptomes attached to microparticles

10X Chromium (3’)



Sequencing



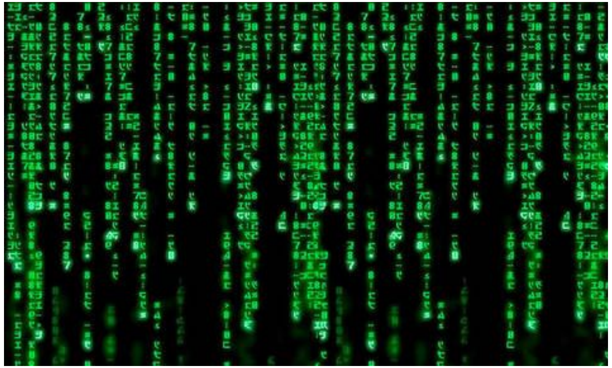


From

Nucleotide sequences

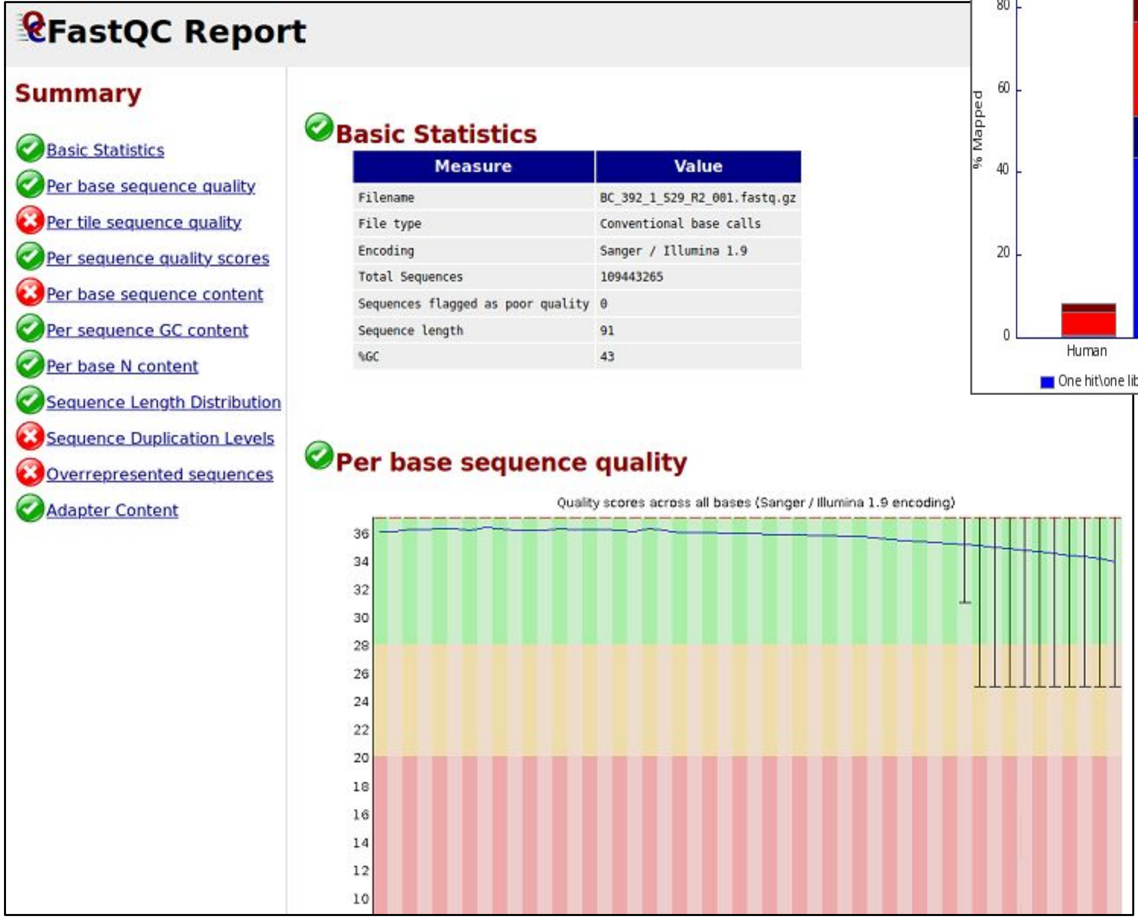
To a

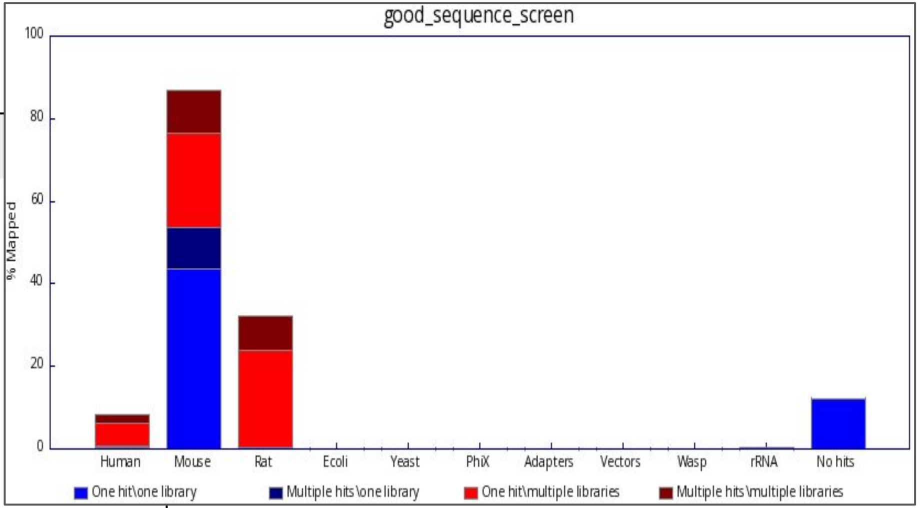
Raw count matrix




CERTIFIED
BIOINFORMAGICIAN

Reads QC

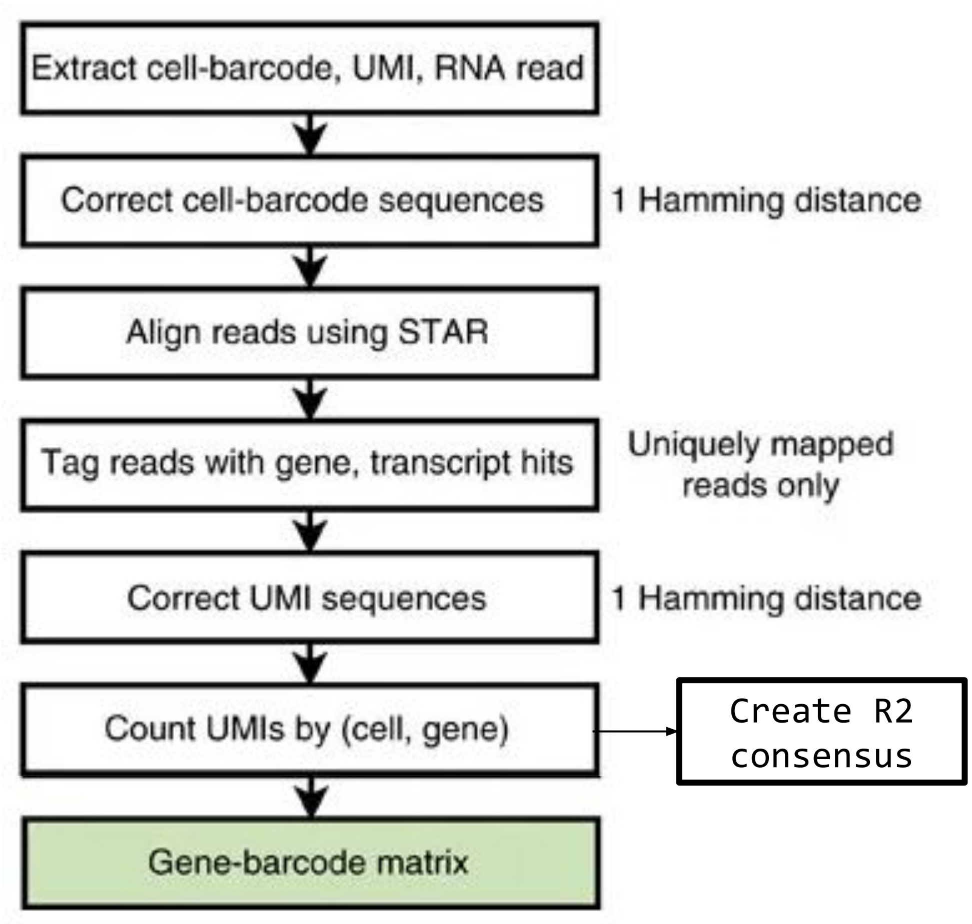




- As usual : FASTQC, FastqScreen, …
- 10x :
  - Special attention to R1 : cell barcode + UMI (no N)
  - Control of the 4 sample libraries

Reads processing workflows

Mapping-based (STAR)



Pseudomapping-based (kallisto bustools)

kallisto index
Generate species index

↓

kallisto bus
Pseudoalign single cell reads

↓

bustools correct
Correct hamming distance 1 barcodes

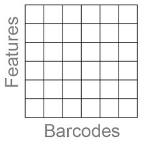
↓

bustools sort
Organize BUS file according to barcode, UMI, set, and multiplicity

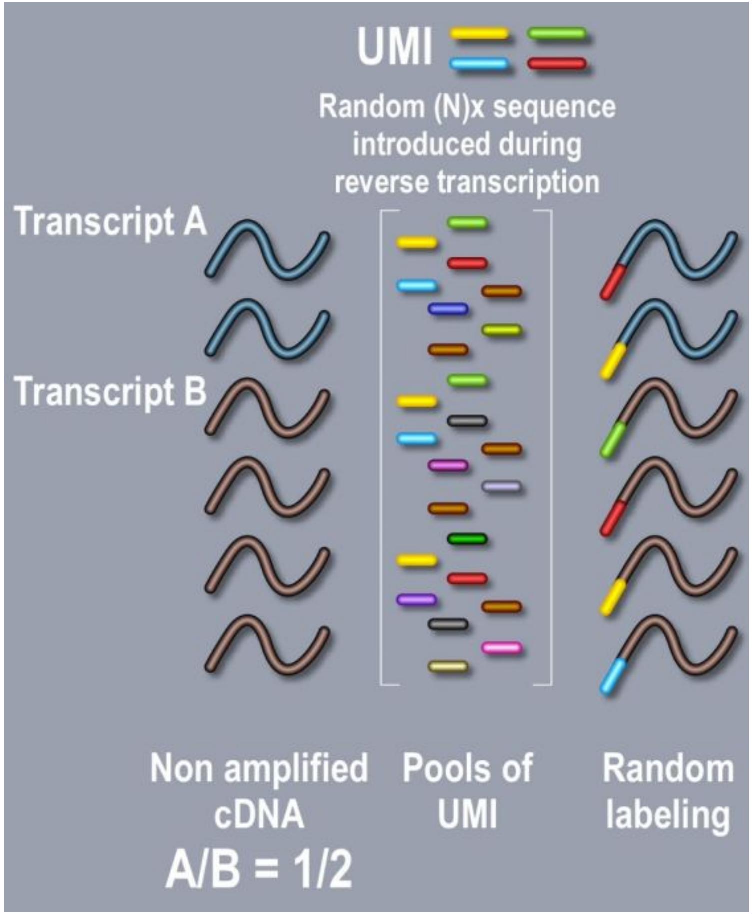
↓

bustools count
Generate UMI count matrix

↓

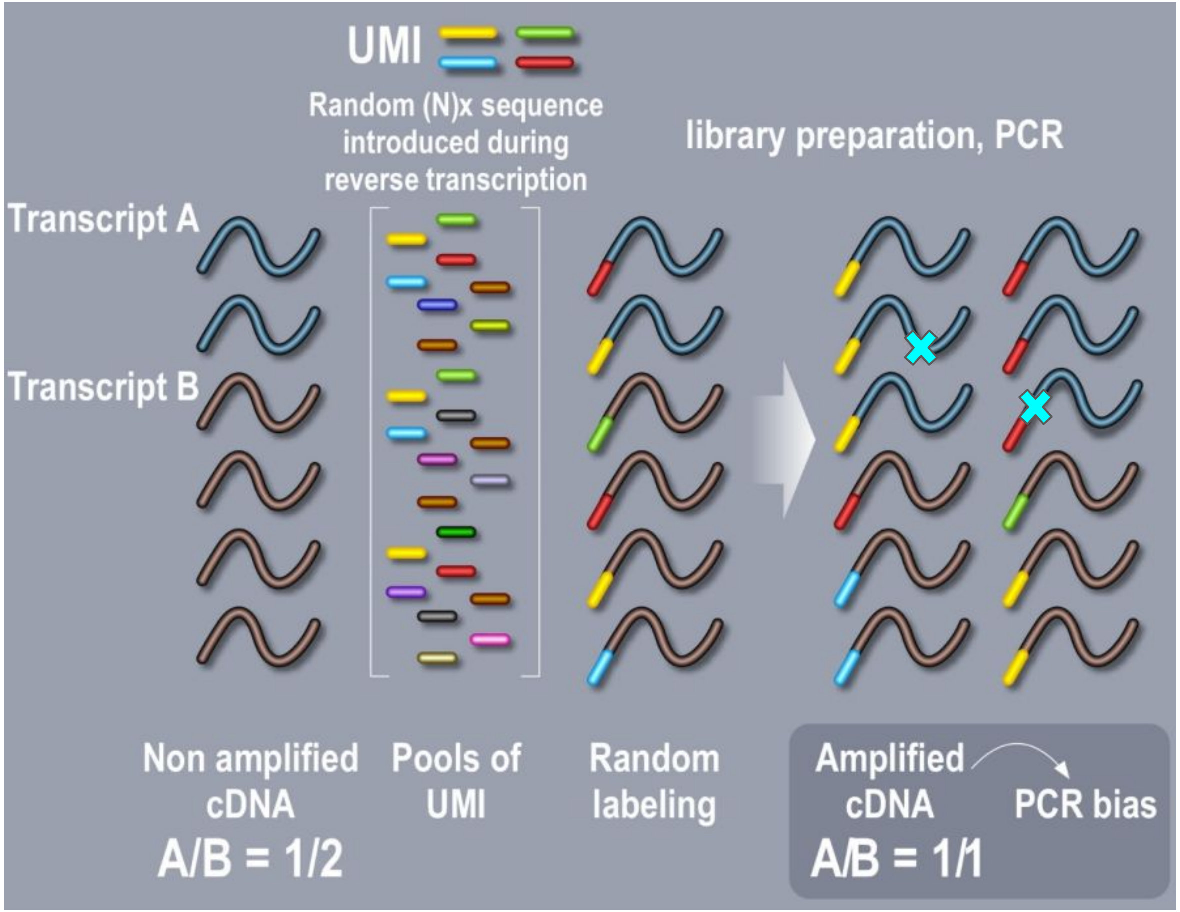


Focus on : Unique Molecule Identifiers



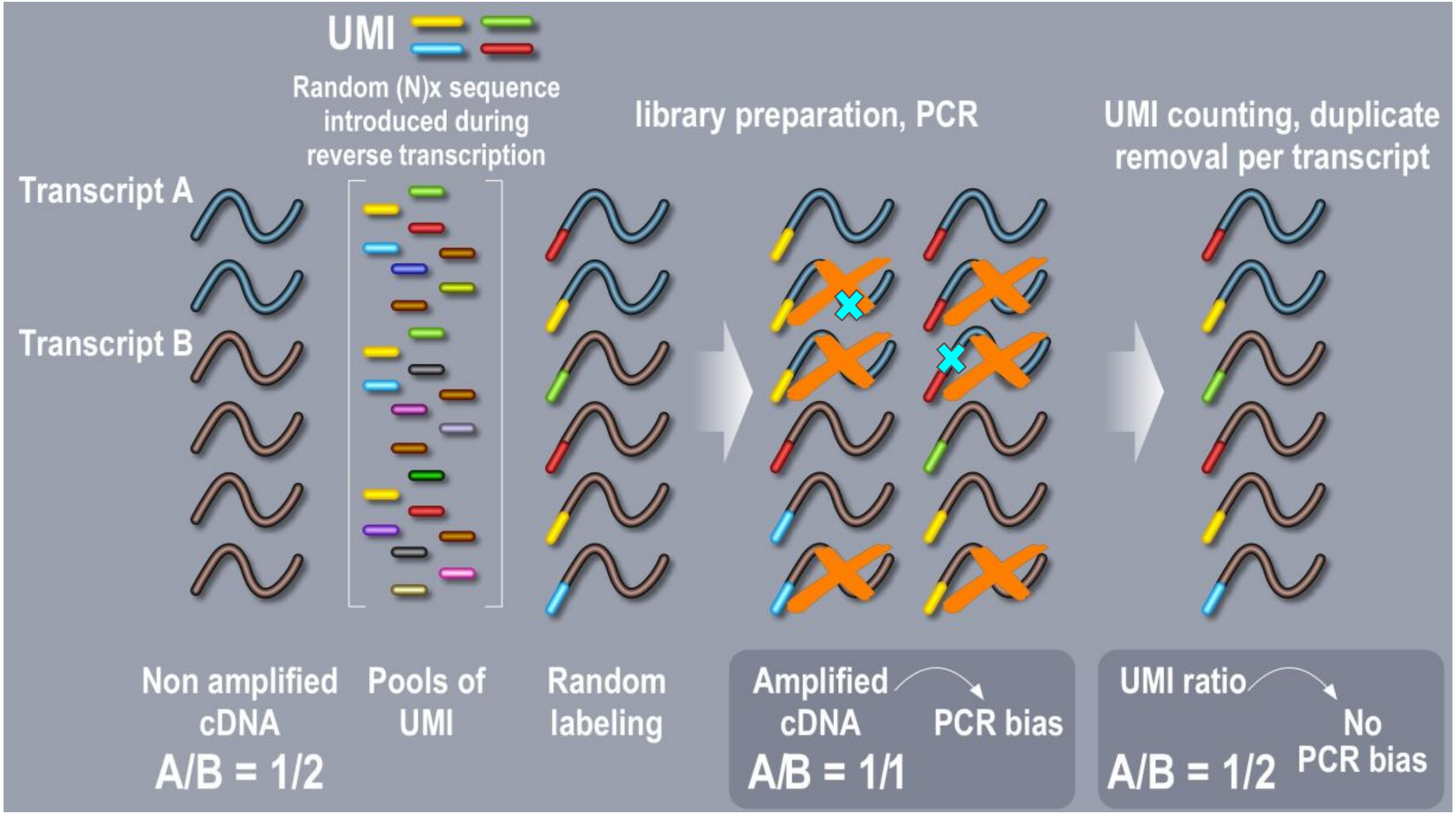
Islam et al., Nature Methods (2014)
Sludy by Agnès Paquet

Focus on : Unique Molecule Identifiers



Islam et al., Nature Methods (2014)
Sludy by Agnès Paquet

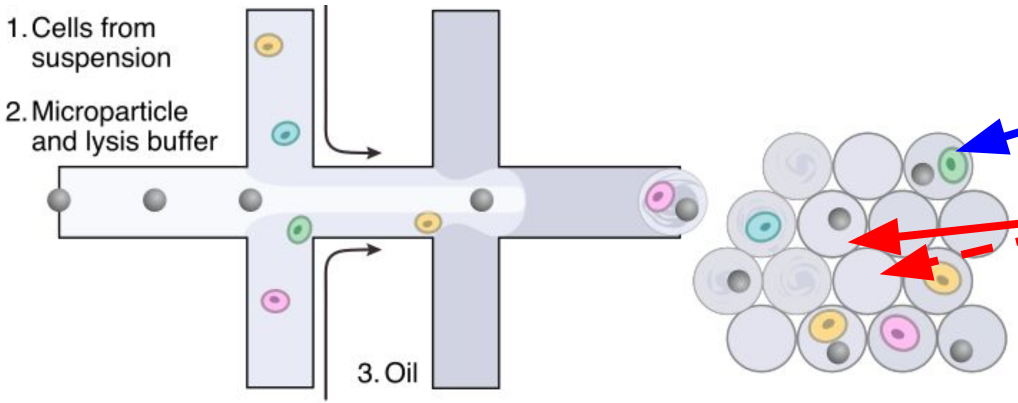
Focus on : Unique Molecule Identifiers



Islam et al., Nature Methods (2014)

Sludy by Agnès Paquet

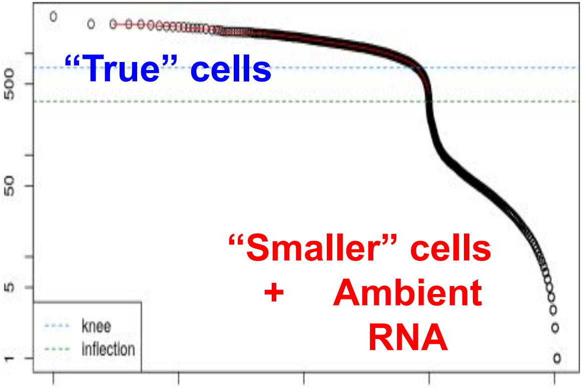
Focus on : Empty droplets filtering



THERE IS RNA HERE (CELL IN GEM)

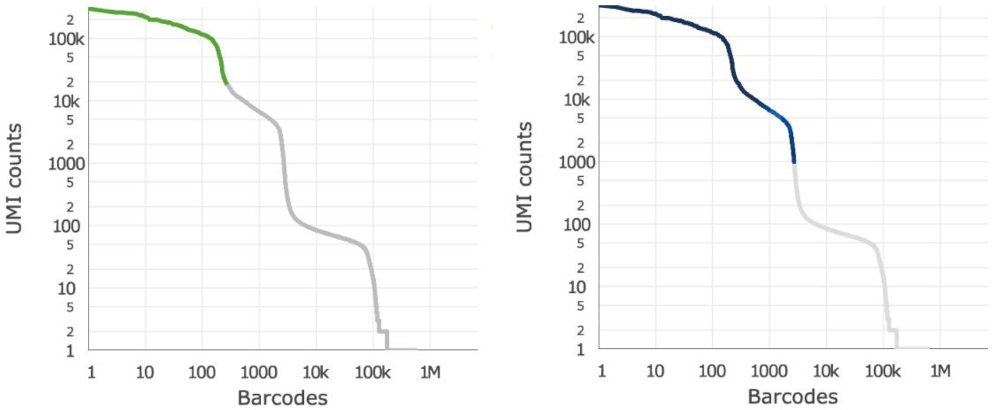
THERE IS RNA HERE TOO ! (NO CELL = AMBIENT)

Kneeplot



<<

EmptyDrops



Lun et al. Genome Biology (2019)

From a

Raw count matrix

To a

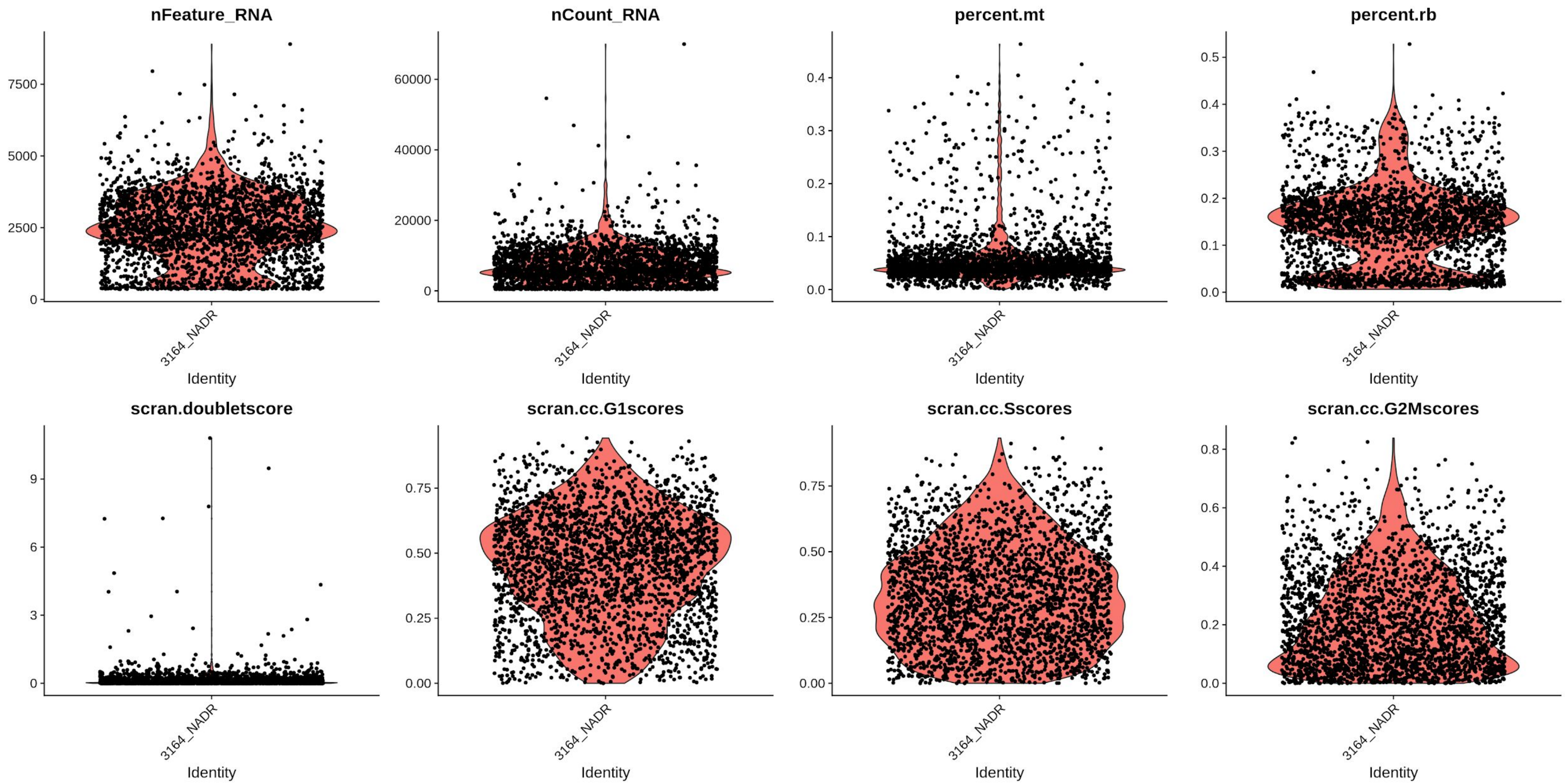
Normalized matrix

Cell QC considerations

- The number of unique genes detected in each cell :
    - Low-quality cells or empty droplets will often have very few genes
    - Cell doublets or multiplets may exhibit an aberrantly high gene count
- Similarly, the total number of molecules detected within a cell (correlates strongly with unique genes)
- The percentage of reads that map to the mitochondrial genome :
    - Low-quality / dying cells often exhibit extensive mitochondrial contamination
- Other QC criteria to measure :
    - Cell cycle phase / score
    - Nuclear riboprotein-coding genes expression

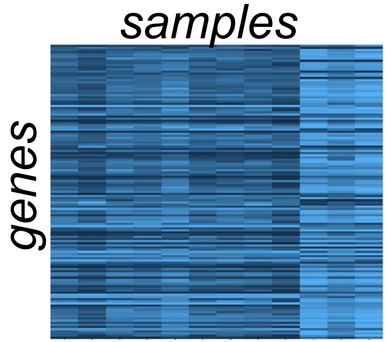
Ilicic et al. Genome Biology (2016)

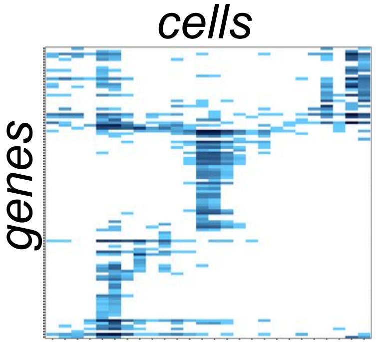
Cell QC : metrics



Matrix normalization : *Houston, we have a problem...*

	BULK	SINGLE-CELL
Total RNA	100 ng (~10.000 cells)	10 pg (per cell)
mRNA	~ 5 ng (~10.000 cells)	<< 1 pg (per cell)
Reads	~100 million	~ 50 k (per cell)





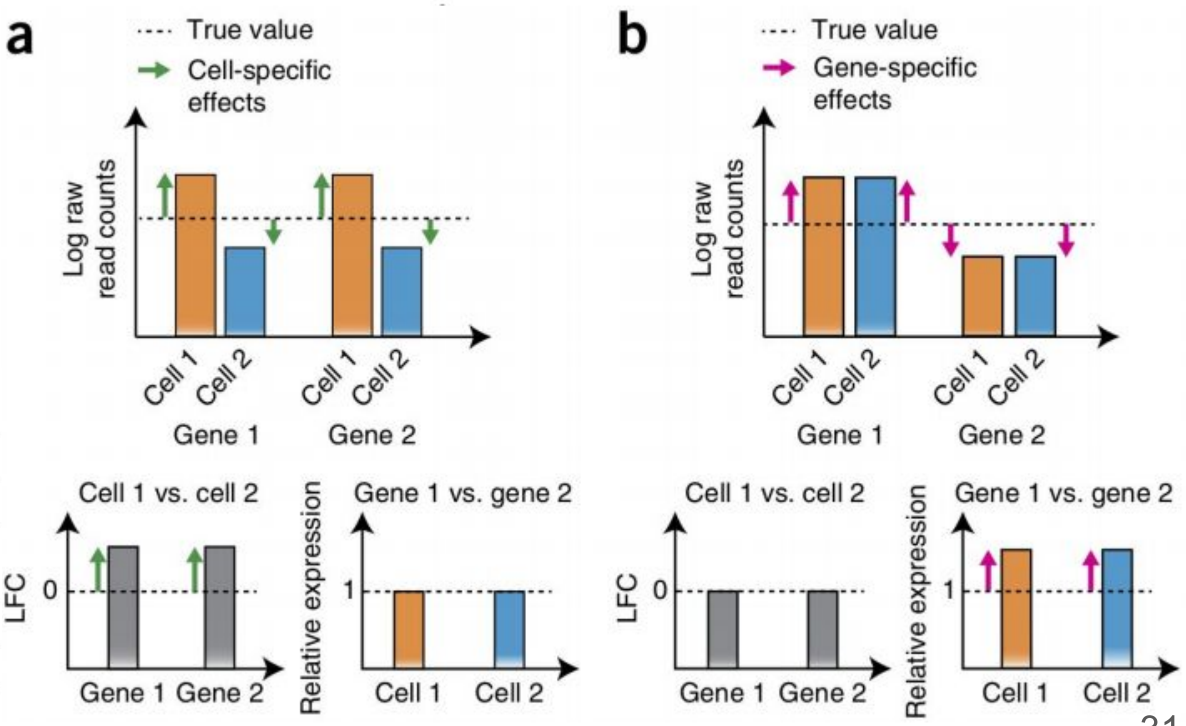
SC MATRIX IS SPARSE ! (ie, mostly filled with zeros)

Matrix normalization : different levels

- Process of identifying and removing systematic variation not due to real differences between RNA treatments i.e. differential gene expression.

- Cell-specific effects

- Gene-specific effects



c

	Cell-specific effects	Gene-specific effects	Not removed by UMIs
Sequencing depth	✓		✓
Amplification	✓	✓	
Capture and RT efficiency	✓	✓	✓
Gene length		✓	
GC content	✓	✓	✓
mRNA content	✓		✓

Bulk normalization methods are KO

- **RPKM/FPKM** (Reads/Fragments per kilobase of transcript per million reads of library) : Normalize for sequencing depth and transcript length at the same time => KO **if you DO NOT have full-length data**

- **Global scaling** (eg: Upper Quartile) : KO **if you have too many zeros**

- **Size factors** calculation (eg: Estimation of library sampling depth) :
  - DESeq2, edgeR suppose that >=50% of genes are **NOT DE**
  - KO **if you have too many zeros**

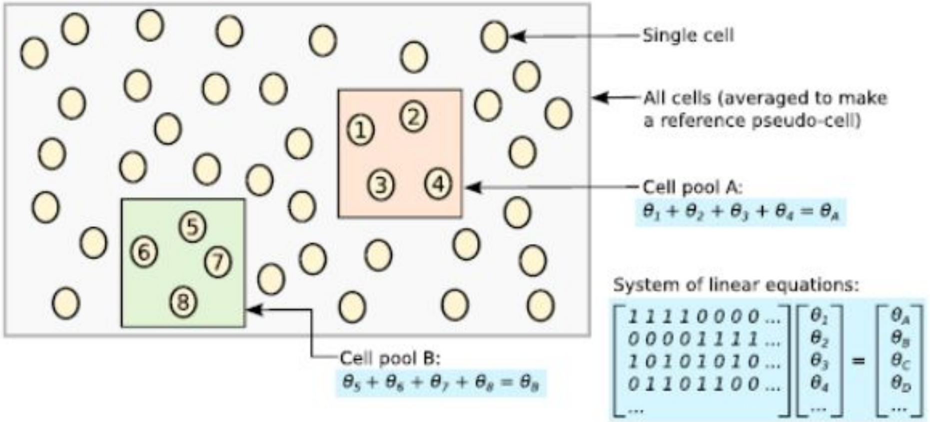
- **TPM/CPM** : KO if a **small number of genes** carry **most of the signal**

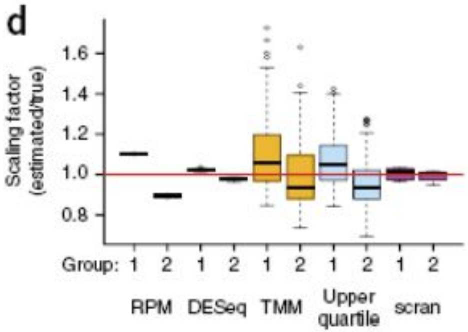
=> Rough solution : global log-normalization / Z-scoring

Paquet 2018

Matrix normalization : scaling by factors

- Alternative method to compute the size factors
- Pool cells to reduce the number of zeros
- Estimate the size factors for the pool
- Repeat many time and use deconvolution to estimate each cell size factor
- Implemented in **scater/scran** packages



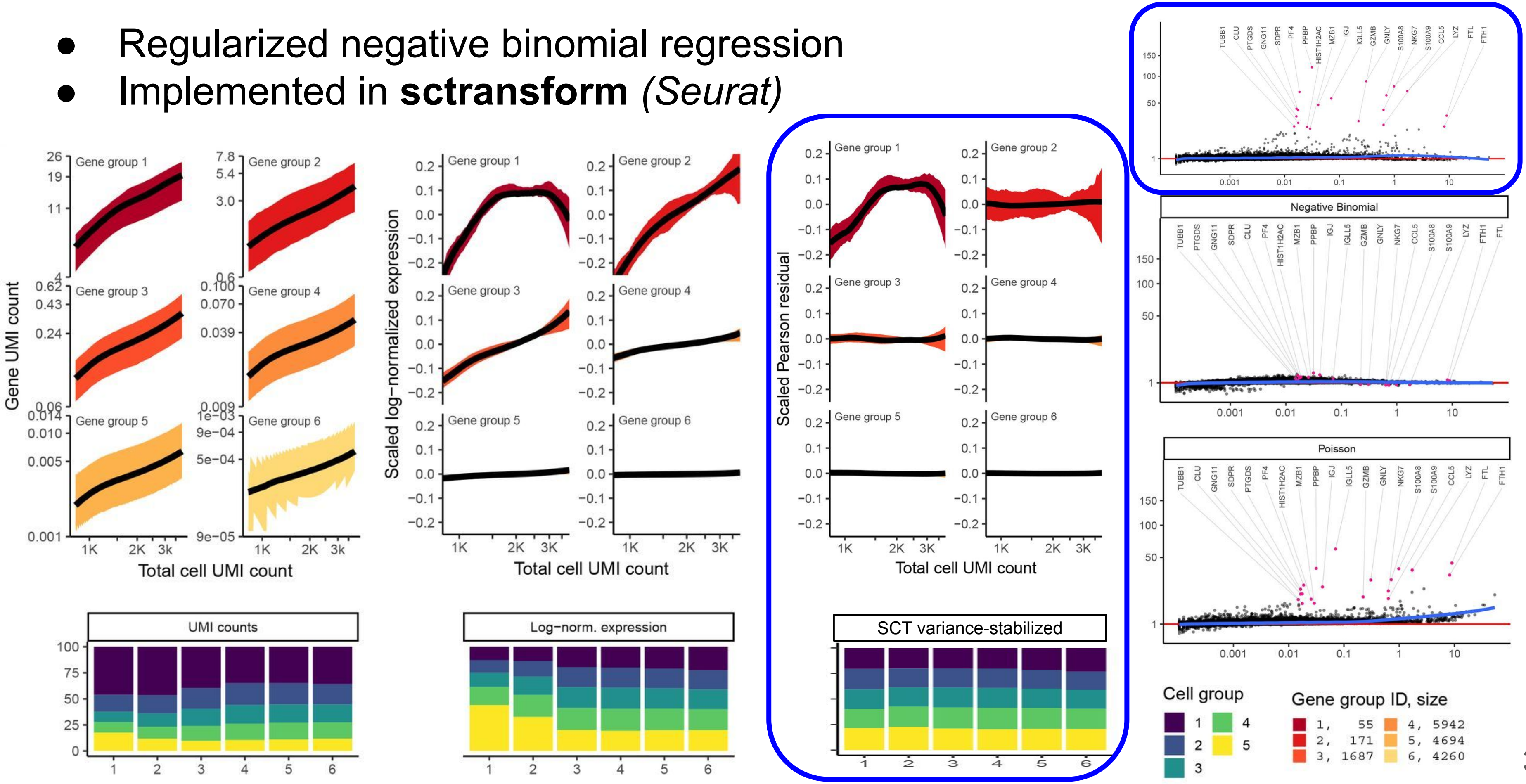


Vallejos C, 2017

Lun et al (2018)

Matrix normalization : variance stabilization

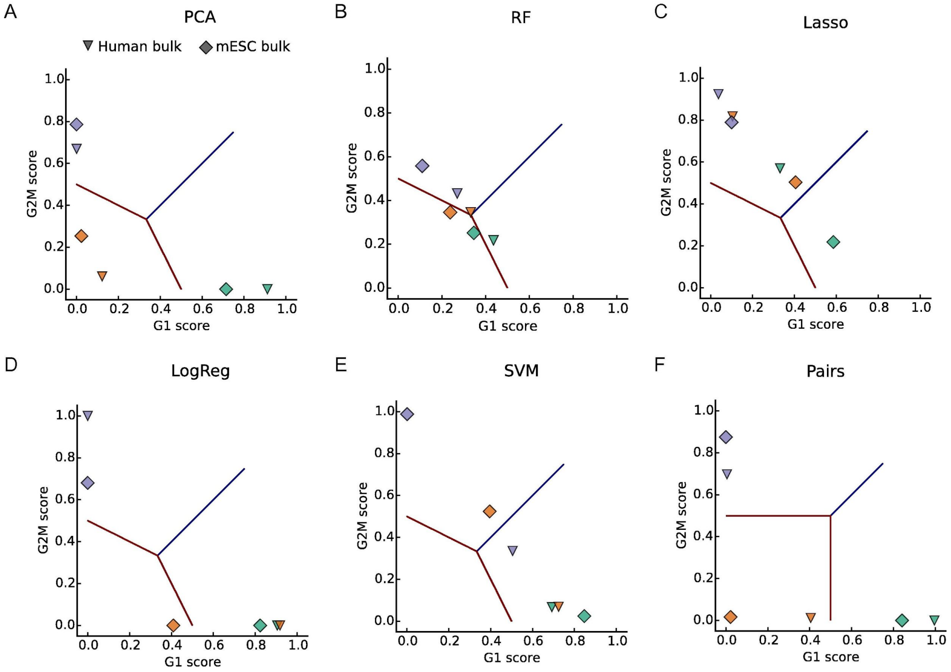
- Regularized negative binomial regression
- Implemented in **sctransform** *(Seurat)*

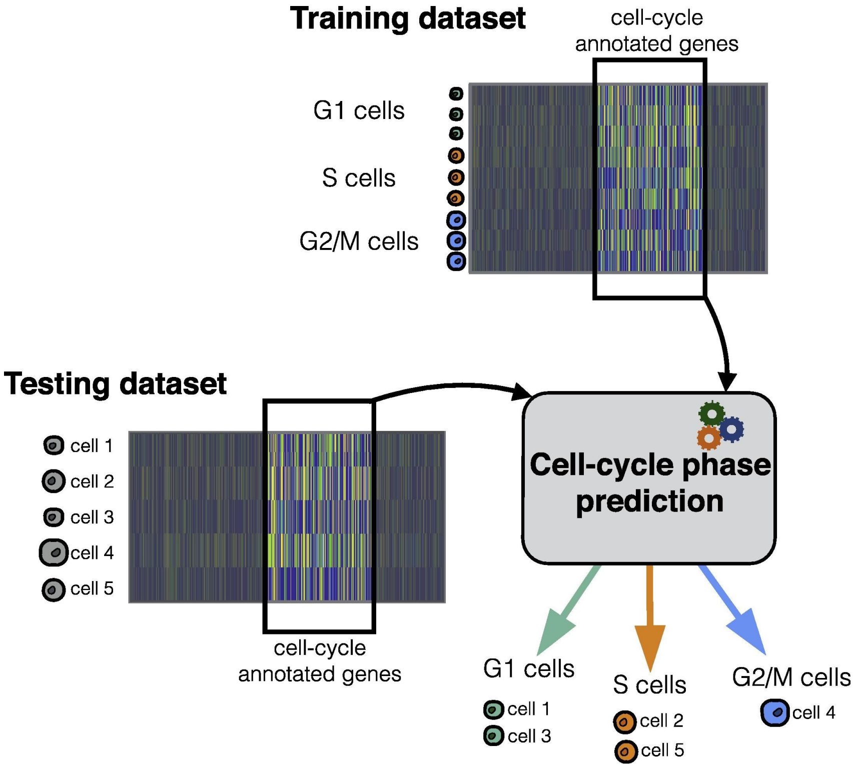


Hafemeister et al, BioRxiv (2019)

Known biases : Cell cycle phase

- Training on reference set with the 3 phases identified
- Use pairs of differential genes
- Apply model pairs to new dataset and assign phases
- Implemented in **cyclone** *(scran)*

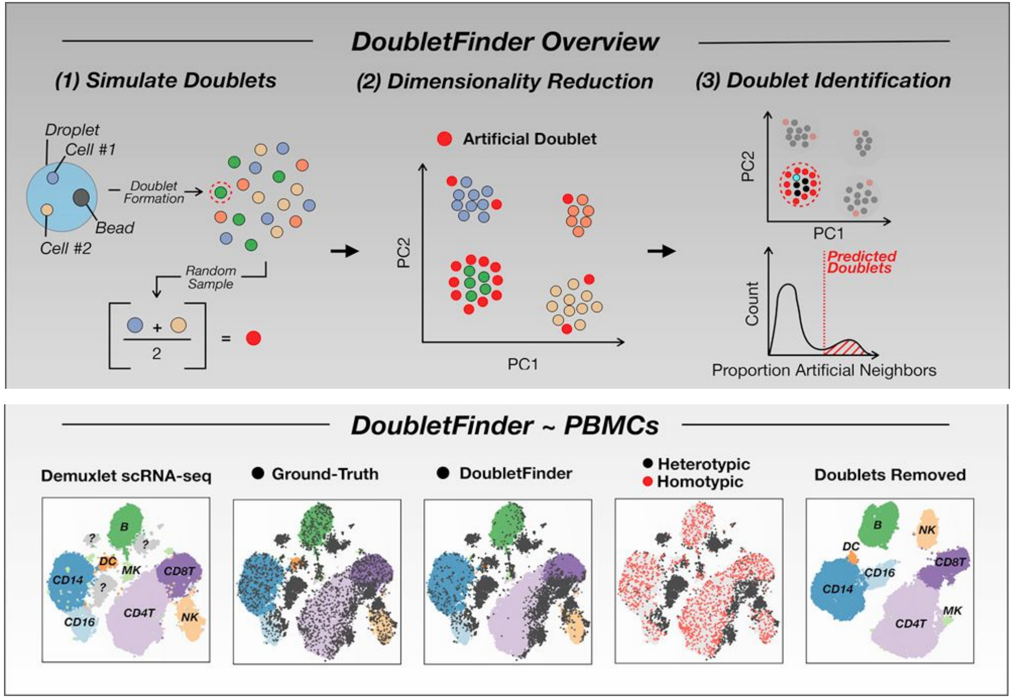




Scialdone et al., Methods (2015)

Known biases : Cell doublets

- Two types of doublets :
  - Cells of the same type => higher global expression
  - Cells of the different types => artificial hybrid
- Methods : generate random artificial doublets, capture all

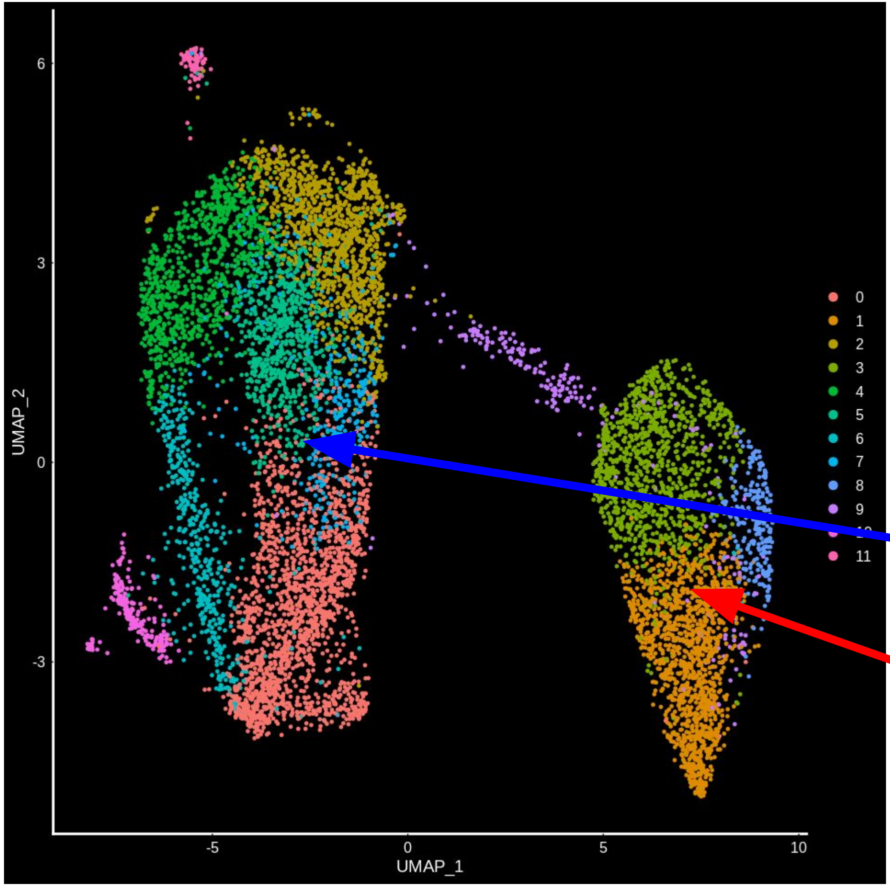


McGuinnis et al., Cell Sys (2019)

	AUC	pAUC90	pAUC95	pAUC97.5	AUPRC
ch_cell-lines					
• libsize	0.60	0.54	0.53	0.52	0.17
• features	0.60	0.55	0.54	0.53	0.19
• dblCells	0.64	0.62	0.61	0.60	0.37
• cxds	0.65	0.59	0.57	0.55	0.26
• dblDetection	0.66	0.66	0.65	0.65	0.44
• scrublet	0.69	0.65	0.64	0.63	0.41
• dblFinder	0.69	0.66	0.65	0.65	0.45
• hybrid	0.70	0.64	0.63	0.61	0.40
• bcds	0.70	0.66	0.64	0.62	0.43
ch_pbmc					
• dblCells	0.63	0.57	0.56	0.54	0.31
• libsize	0.78	0.63	0.57	0.54	0.44
• scrublet	0.78	0.67	0.63	0.59	0.52
• cxds	0.78	0.69	0.65	0.61	0.54
• features	0.79	0.62	0.57	0.54	0.45
• bcds	0.81	0.71	0.66	0.60	0.58
• hybrid	0.82	0.73	0.67	0.62	0.61
• dblDetection	0.82	0.75	0.69	0.62	0.63
• dblFinder	0.84	0.74	0.68	0.62	0.64
demuxlet					
• dblCells	0.79	0.70	0.65	0.60	0.46
• libsize	0.81	0.58	0.55	0.53	0.30
• features	0.85	0.62	0.57	0.55	0.37
• scrublet	0.87	0.74	0.68	0.62	0.53
• cxds	0.89	0.71	0.63	0.57	0.49
• hybrid	0.91	0.78	0.68	0.58	0.57
• dblDetection	0.91	0.79	0.69	0.58	0.57
• bcds	0.91	0.79	0.71	0.62	0.61
• dblFinder	0.92	0.79	0.70	0.63	0.62
hg-mm					
• libsize	0.87	0.66	0.59	0.54	0.27
• features	0.89	0.68	0.60	0.55	0.30
• dblCells	0.93	0.88	0.84	0.79	0.73
• bcds	0.96	0.87	0.80	0.71	0.64
• hybrid	0.98	0.94	0.90	0.87	0.88
• scrublet	0.99	0.96	0.94	0.91	0.91
• cxds	0.99	0.98	0.98	0.97	0.97
• dblDetection	0.99	0.99	0.98	0.98	0.97
• dblFinder	1.00	0.99	0.99	0.99	0.99

Bais et al., BioRxiv (2019)

Known biases : an IRL example

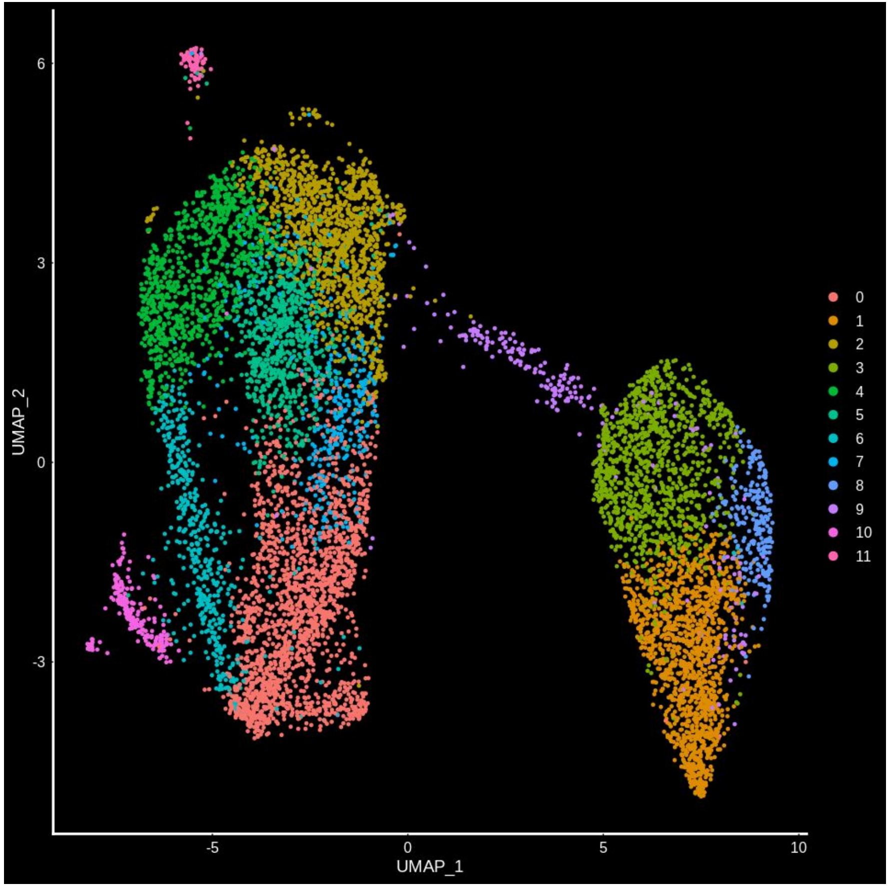


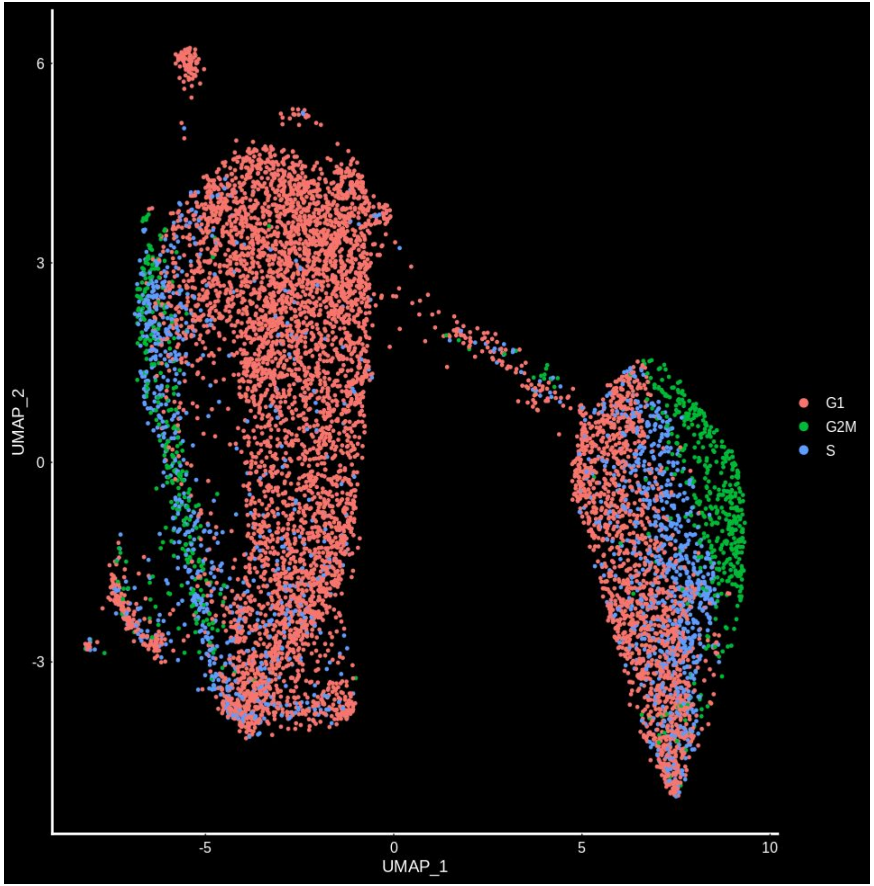
- 10X 3’ scRNAseq v2
- Osteosarcoma metastasis
- 8911 cells x 18613 genes
- PCA (109 PCs retained)
- Louvain clustering
  - 12 clusters
- uMAP representation

Osteoblasts

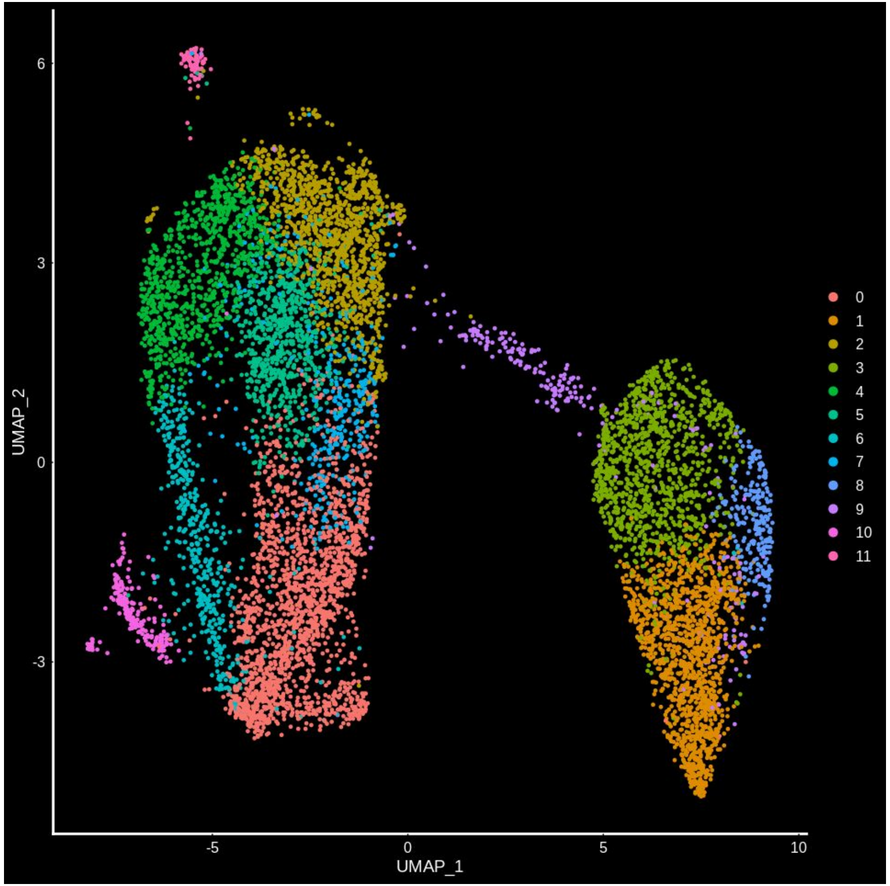
Osteoclasts

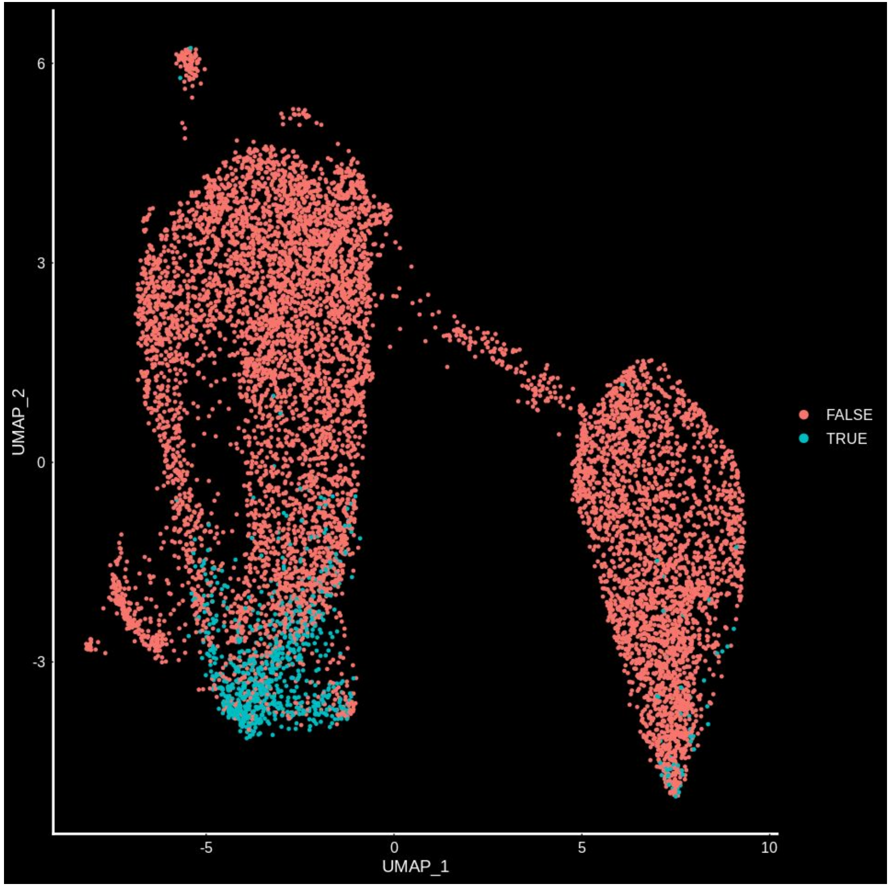
Bias : Cell cycle phases / scores



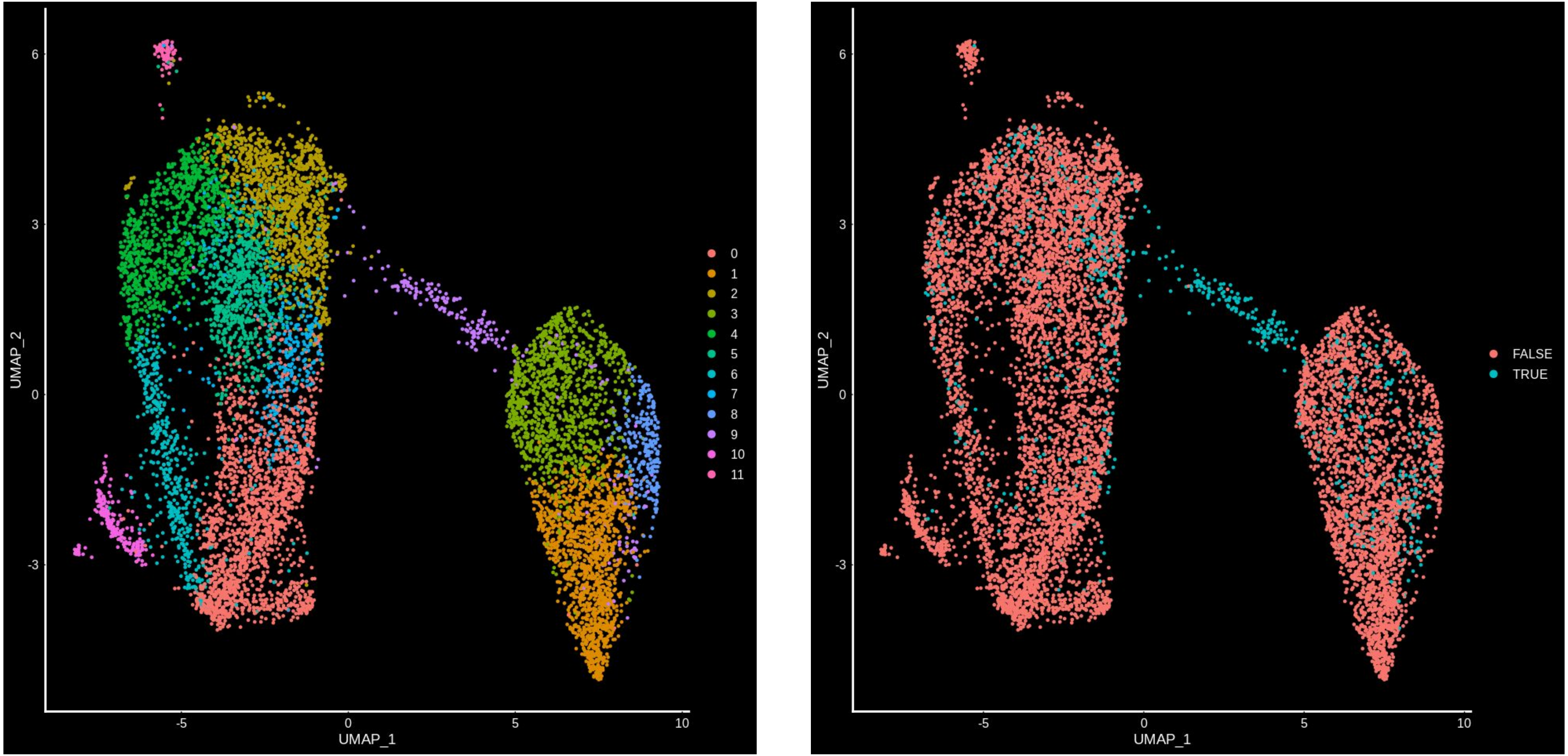


Bias : Dying cells status / score



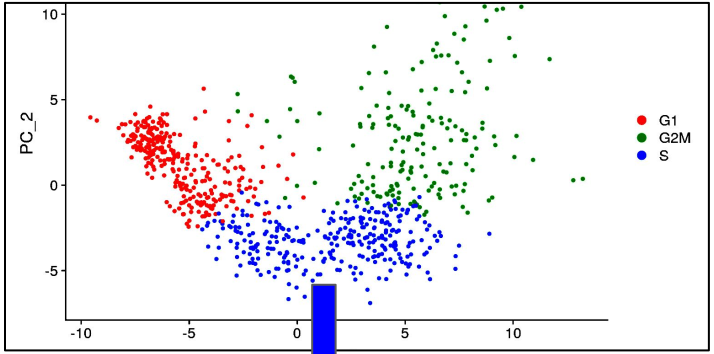


Bias : Cell doublet status / score

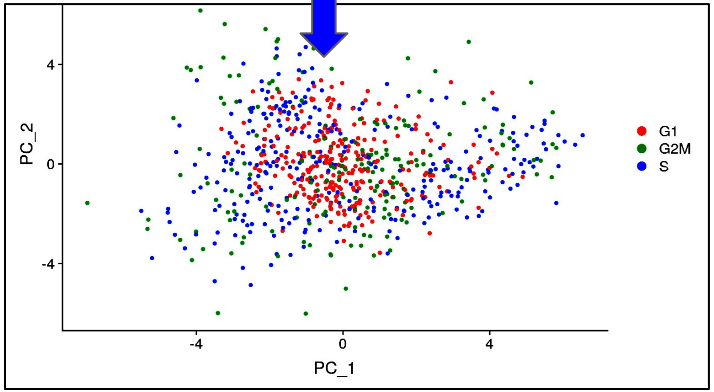


Bias normalization : regression / deblocking

Ex : Cell cycle score

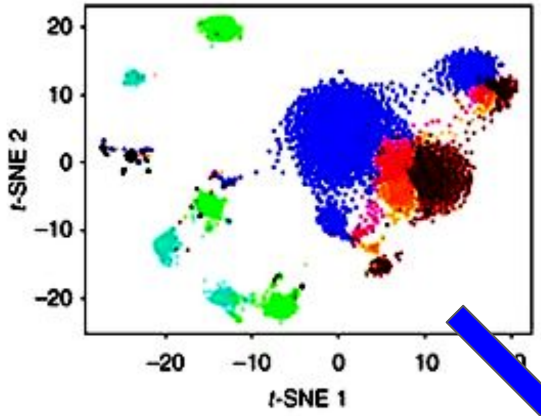


Regression

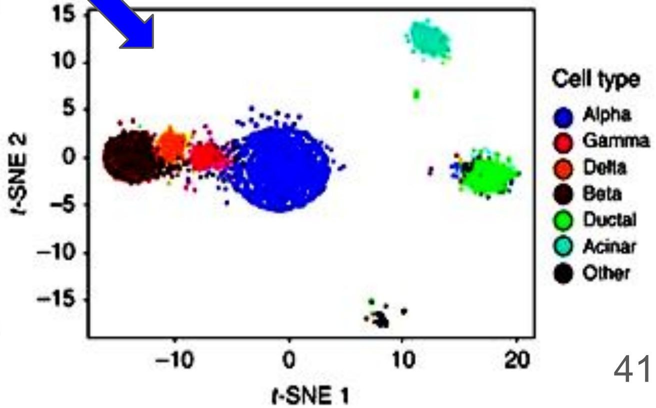


Seurat tutorial

Ex : Batch effect



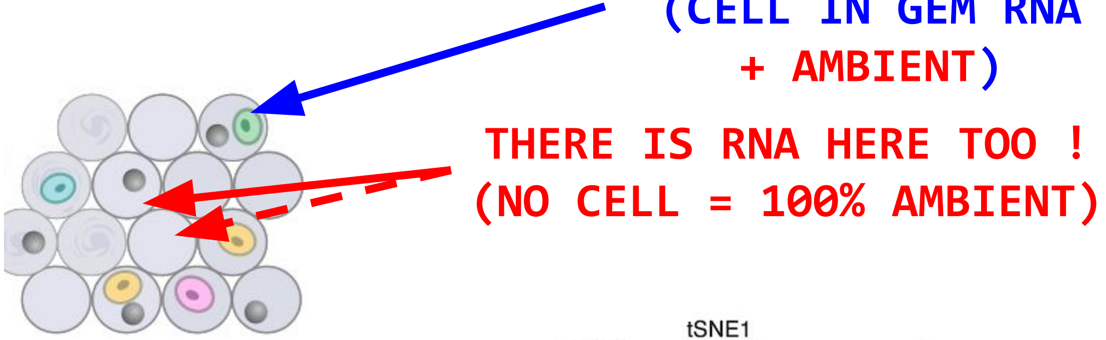
Deblocking

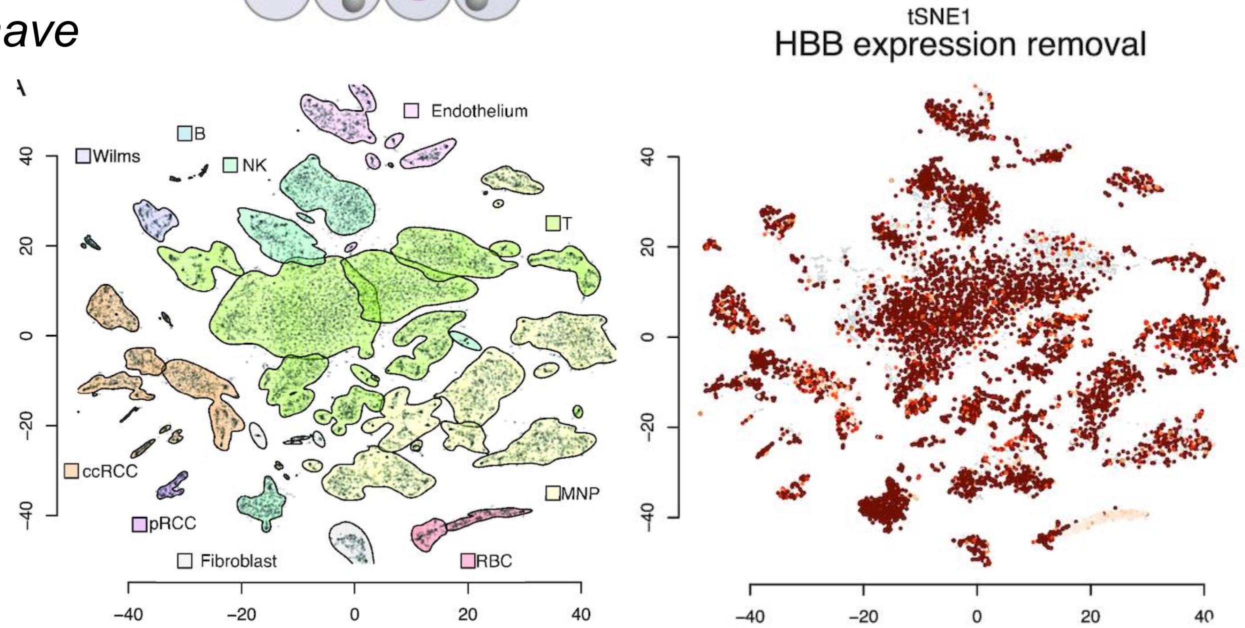


Haghverdi et al. Nature Biotech (2018) (MNN)

Ambient RNA filtering (soupX)

- emptyDrops : removed empty droplets (contained only ambient RNA)
- ***BUT*** *non-empty droplets* ***ALSO*** *have ambient RNA !*
- ***soupX*** determines the amount of ambient RNA in counts, removes it





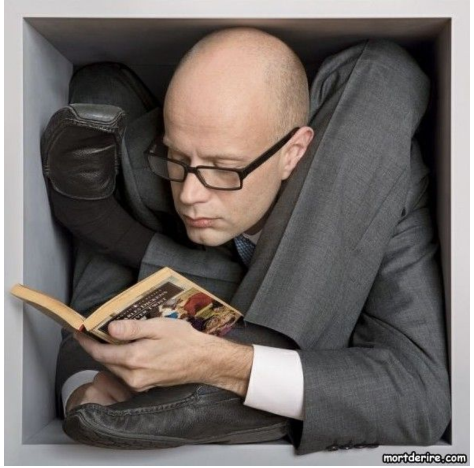
Young et al. GigaScience (2020)

From a

Normalized matrix

To a

Reduced space



Feature selection : Highly variable genes (HVGs)

Postulate : genes with the highest variability should be the most useful to
1. Assess effect of unwanted sources of variation (cell to cell variation)
2. Quantify true biological differences (population to population variation)

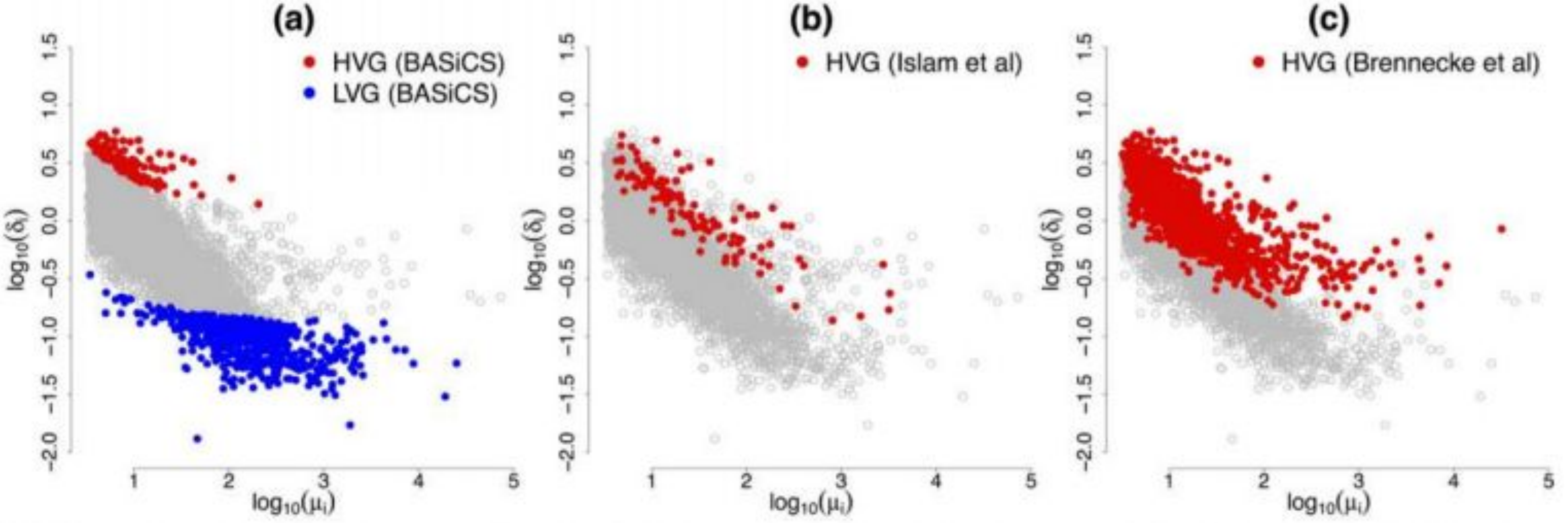


Fig 8. Comparison of HVG detection among different methods. For each of the 7,895 biological genes, posterior medians of biological cell-to-cell heterogeneity term δ_i (log scale) against posterior medians of expression level μ_i (log scale). While the methods described in [16] and [5] only provide a characterisation of HVG, BASiCS is able to detect those genes whose expression rates are stable among cells.

Vallejos et al. Plos Comp Biol (2015)

Dimensionality reduction : simplification + selection

1. Need of an orthogonal space

2. Minimize curse of dimensionality

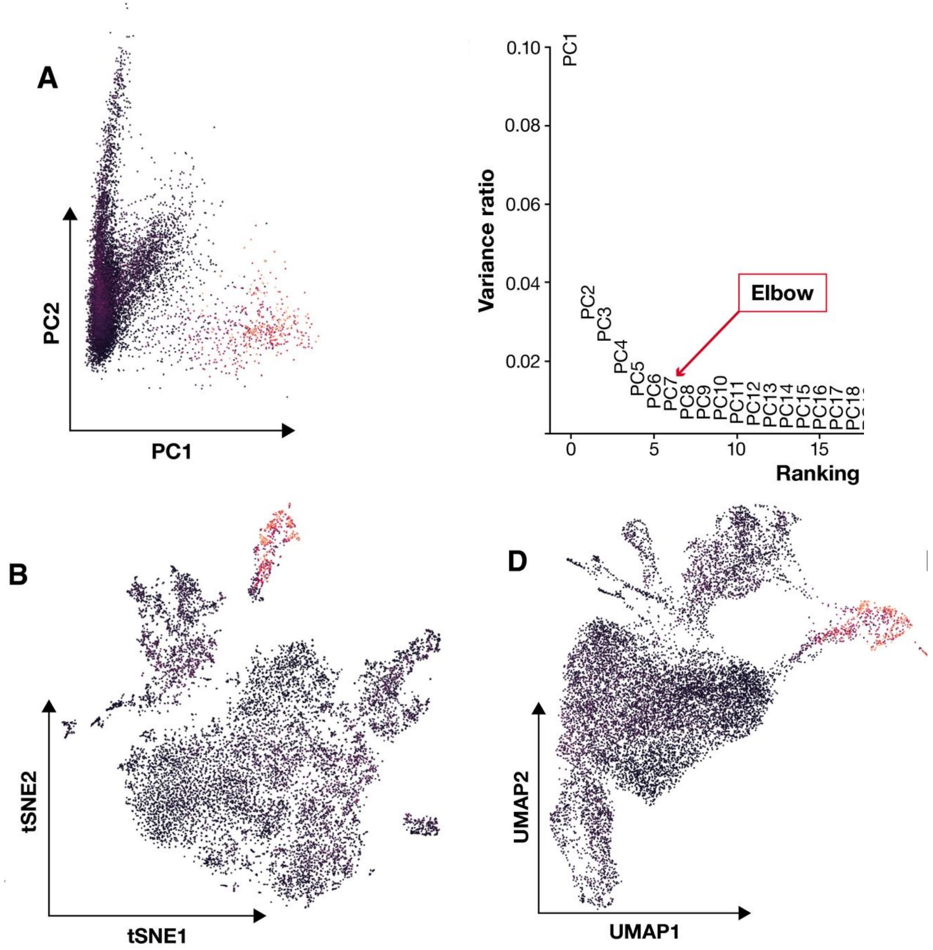
3. Filter out noise

4. Allow visualization

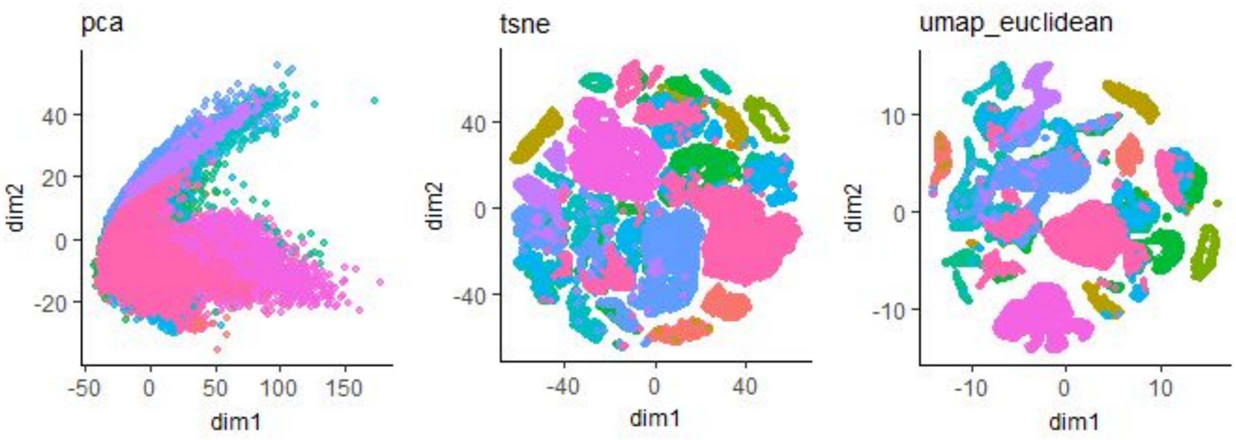
5. Reduce computational load

Popular methods used for single-cell data analysis:

1. PCA
2. ICA
3. tSNE
4. UMAP
5. Others : Diffusion map, Isomap



Dimensionality reduction : PCA / tSNE / uMAP)

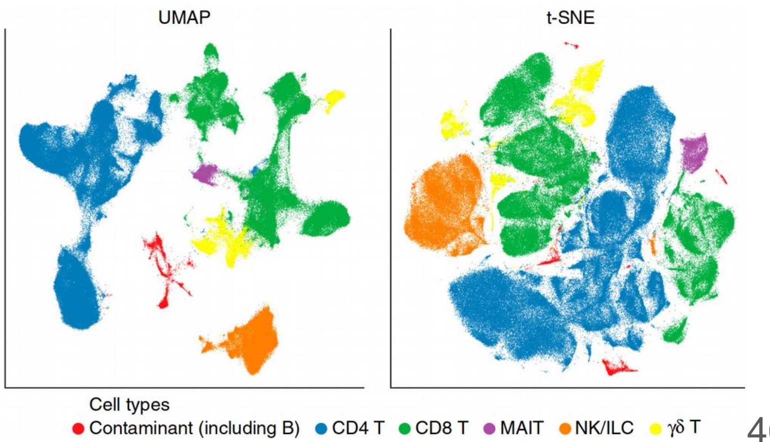


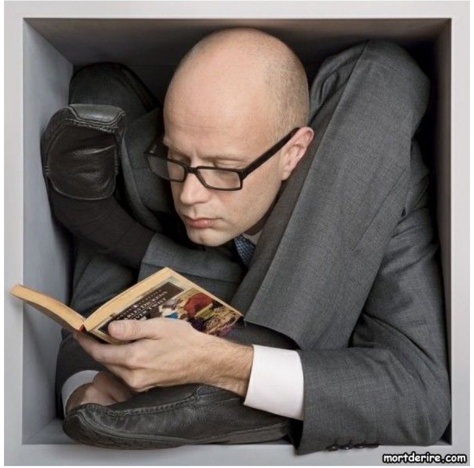
Reduction :

- PCA (on single cell data) is unable to concentrate relationships in 2-3 dimensions only

Visualization : uMAP > tSNE

- Better compaction
- Mostly retains inter-cluster distances
  - Subpopulations
  - Trajectory
- More robust to parameters modification
- (Slightly faster to generate)



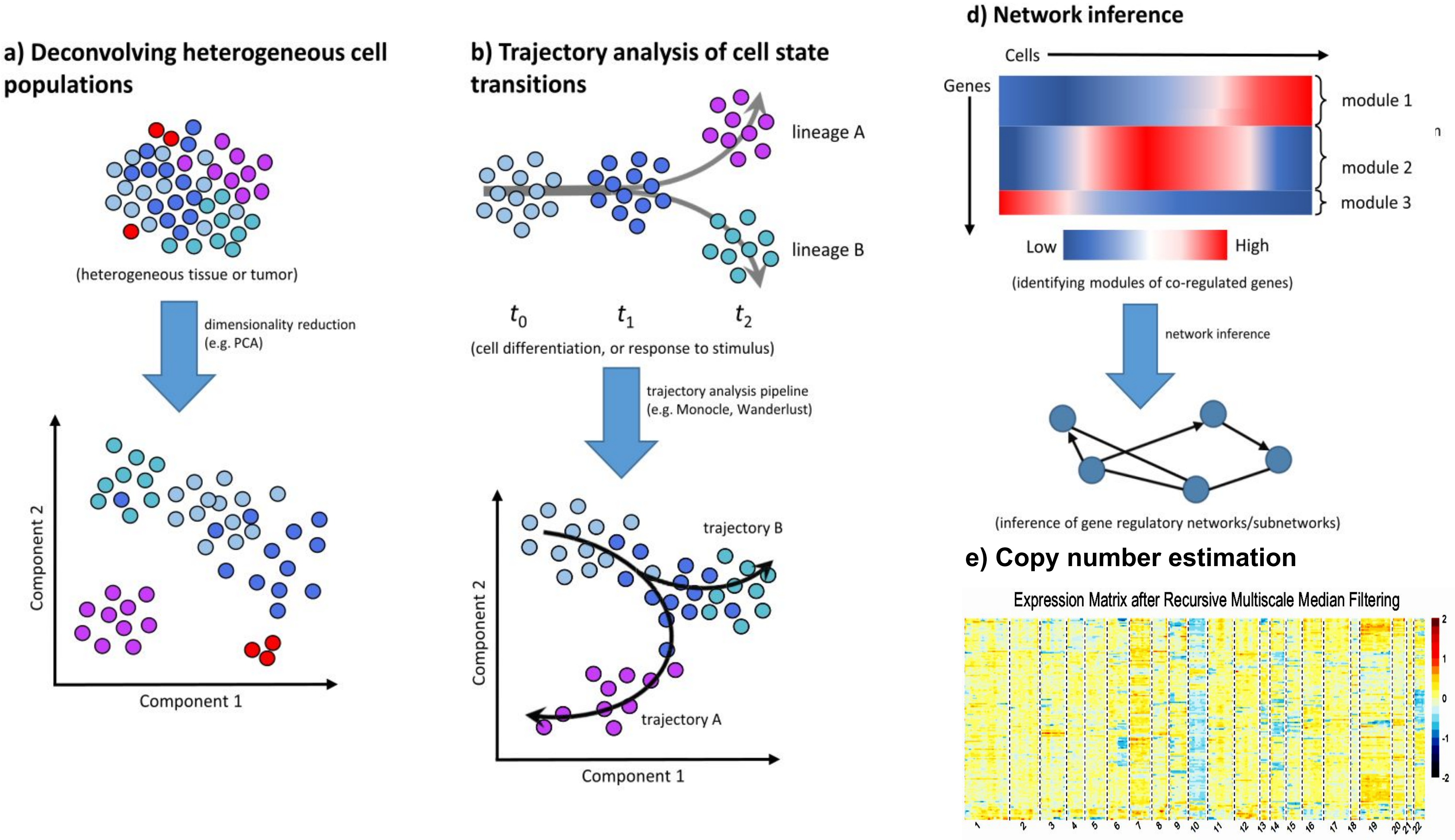


From a

Reduced space

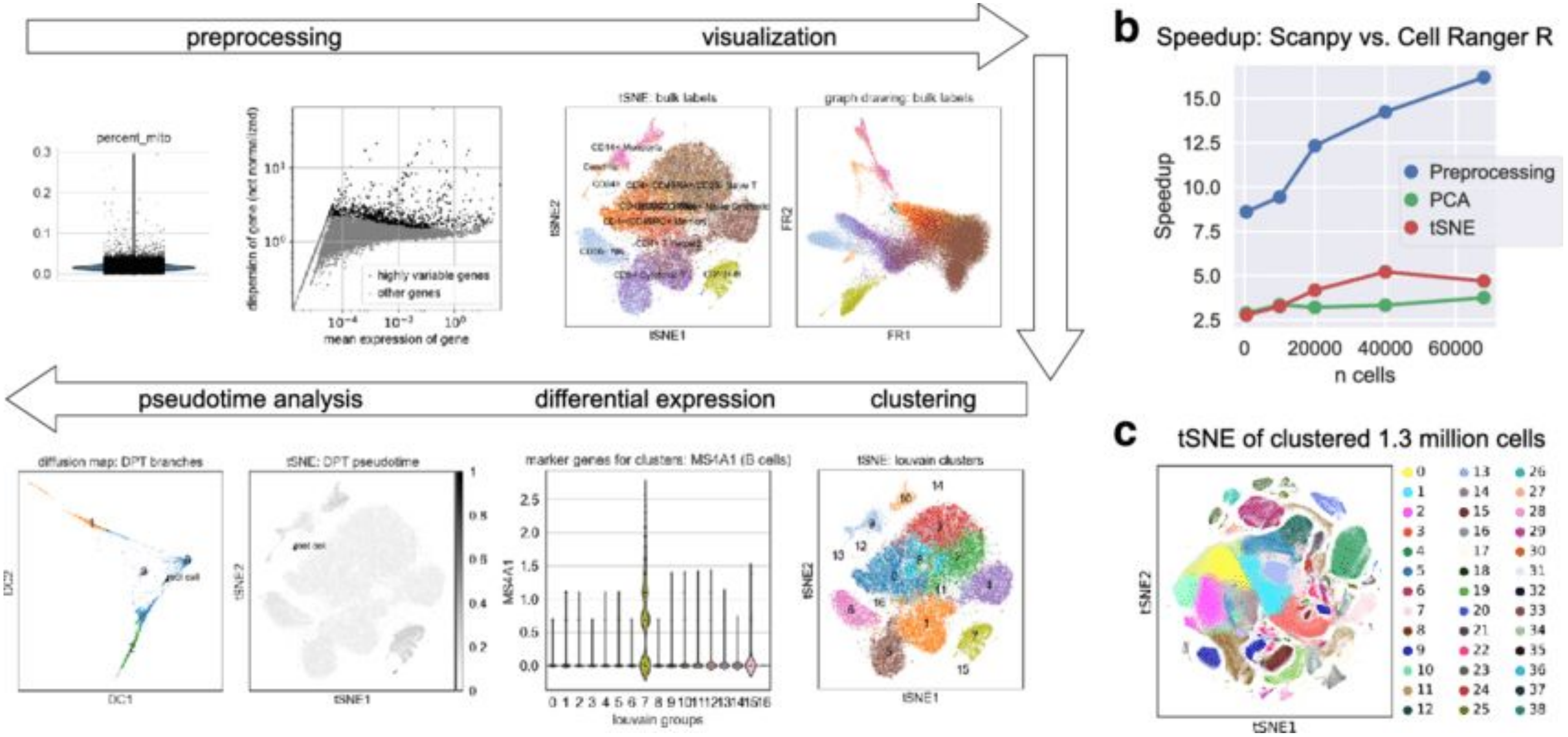
To …

… finally what you wanted !



Liu and Trapnell 2016

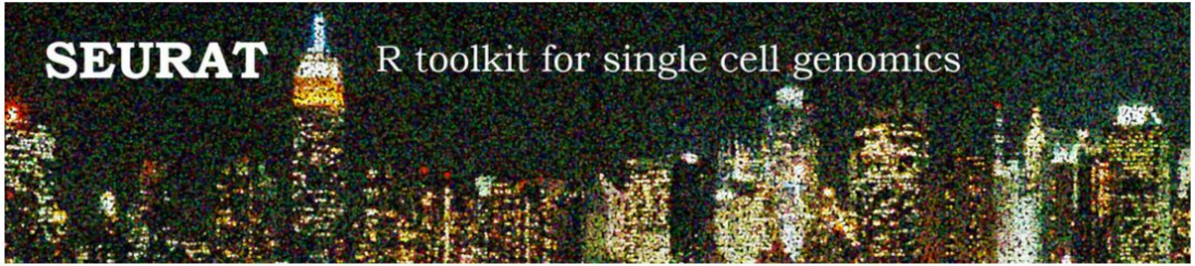
The all-in-one **Python** toolbox : Scanpy



Wolf et al. Genome Biology 2018

The all-in-one **R** toolbox : Seurat

Seurat 4.0.4 | Install | Get started | Vignettes | Extensions | FAQ | News | Reference | Archive



Official release of Seurat 4.0

We are excited to release Seurat v4.0! This update brings the following new features and functionality:

- **Integrative multimodal analysis.** The ability to make simultaneous measurements of multiple data types from the same cell, known as multimodal analysis, represents a new and exciting frontier for single-cell genomics. In Seurat v4, we introduce weighted nearest neighbor (WNN) analysis, an unsupervised strategy to learn the information content of each modality in each cell, and to define cellular state based on a weighted combination of both modalities. In our new paper, we generate a CITE-seq dataset featuring paired measurements of the transcriptome and 228 surface proteins, and leverage WNN to define a multimodal reference of human PBMC. You can use WNN to analyze multimodal data from a variety of technologies, including CITE-seq, ASAP-seq, 10X Genomics ATAC +

Links

Download from CRAN at https://cloud.r-project.org/package=Seurat

Browse source code at https://github.com/satijalab/seurat/

Report a bug at https://github.com/satijalab/seurat/issues

License

GPL-3 | file LICENSE

Community

Code of conduct

Citation

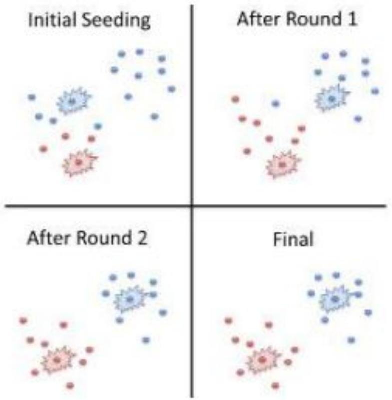
Citing Seurat

Hao*, Hao*, et al., Cell 2021

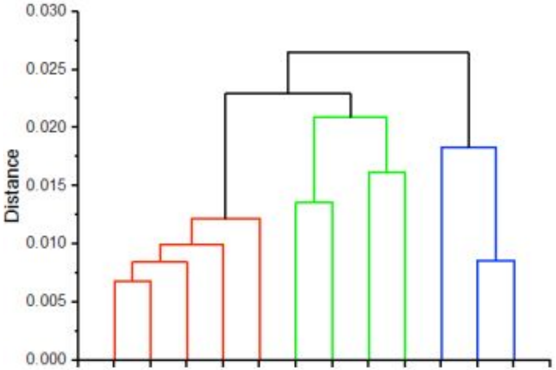
Cell clustering : methods

1) K-means based

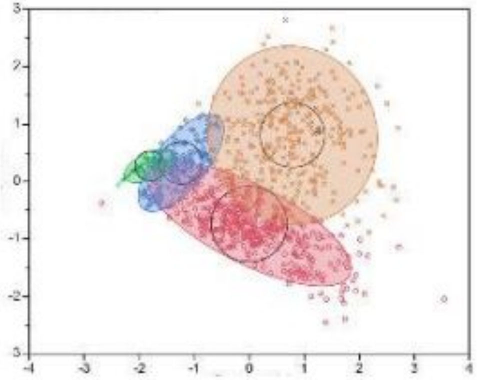
Page et al. BMC Research Notes 2014, 7:829
http://www.biomedcentral.com/1756-0500/7/829



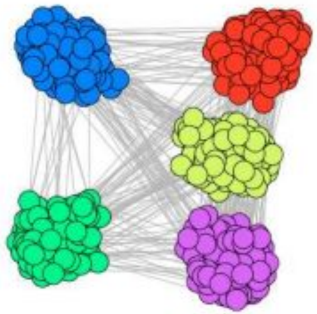
2) Hierarchical clustering



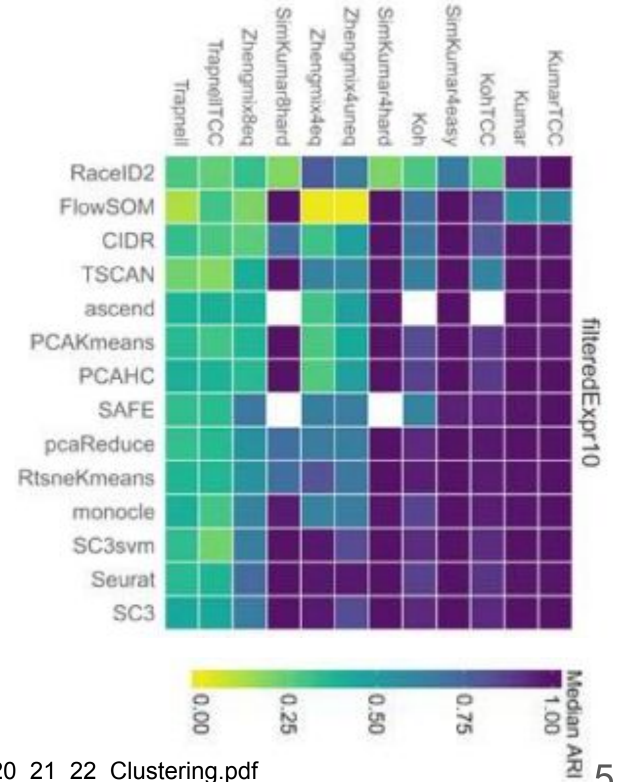
3) Model-based clustering (Mclust)



4) Graph-based clustering

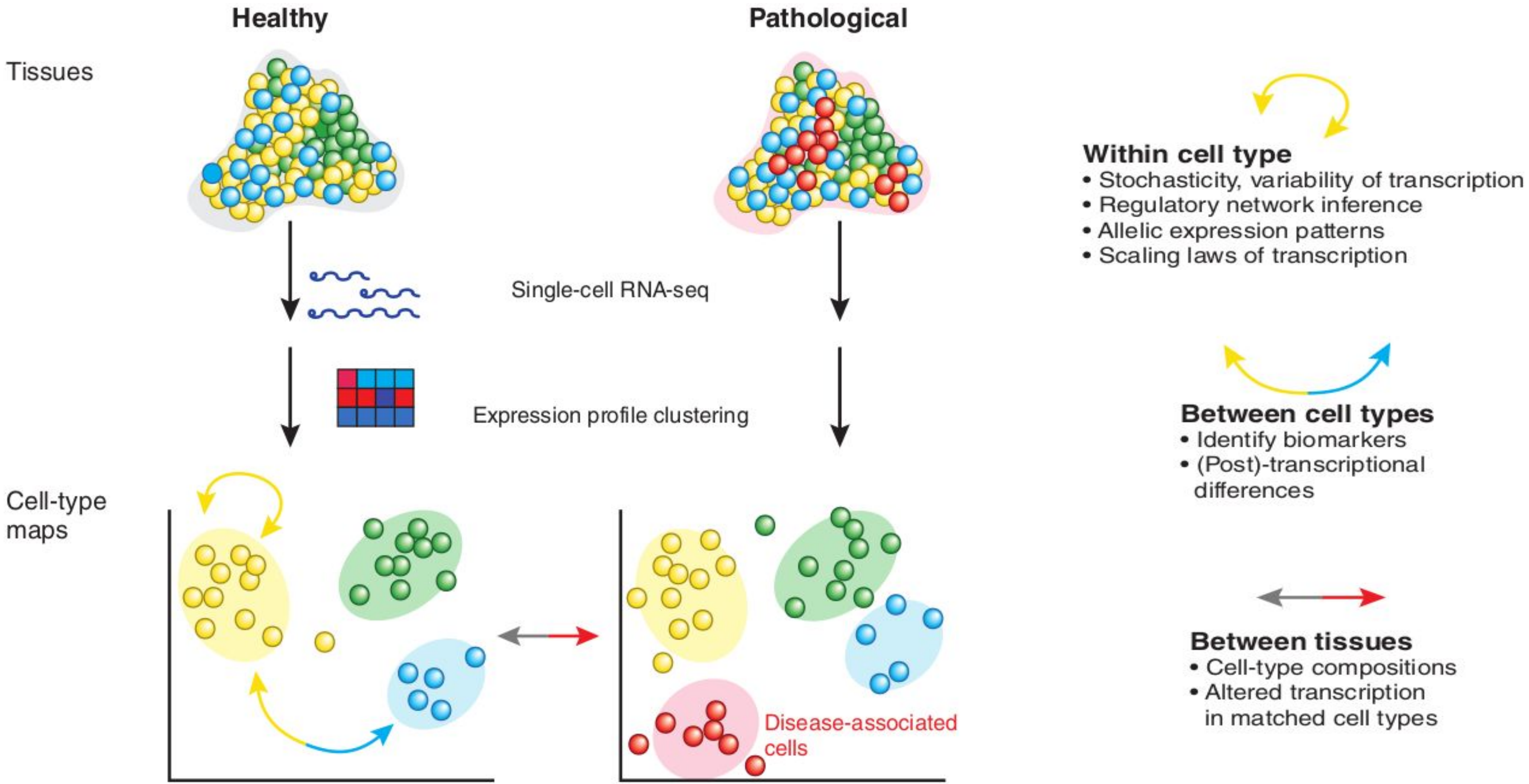


5) Single-cell specific methods

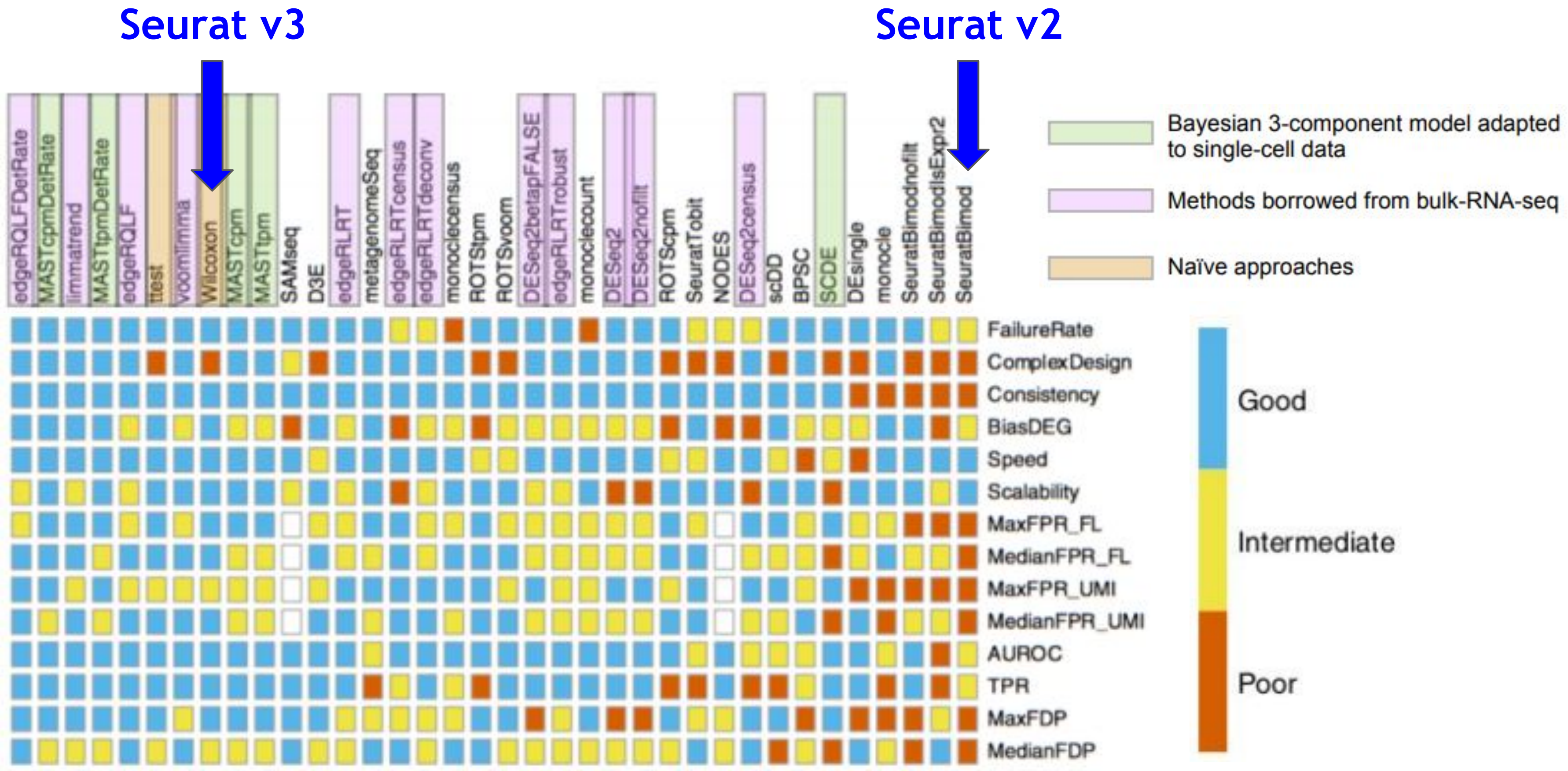


See : https://www.cse.msu.edu/~cse802/S17/slides/Lec_20_21_22_Clustering.pdf

Differential expression analysis

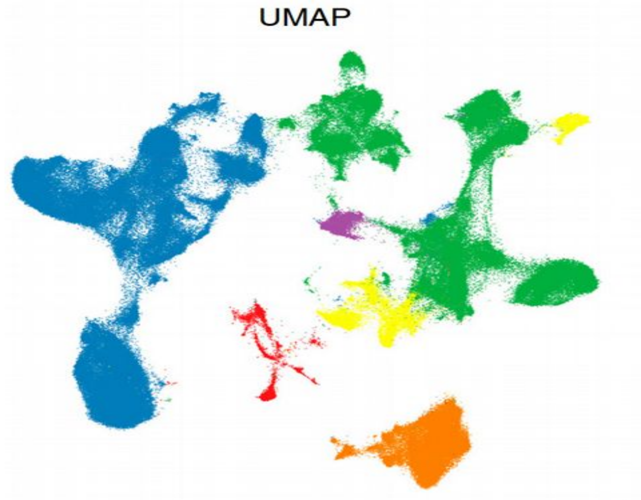


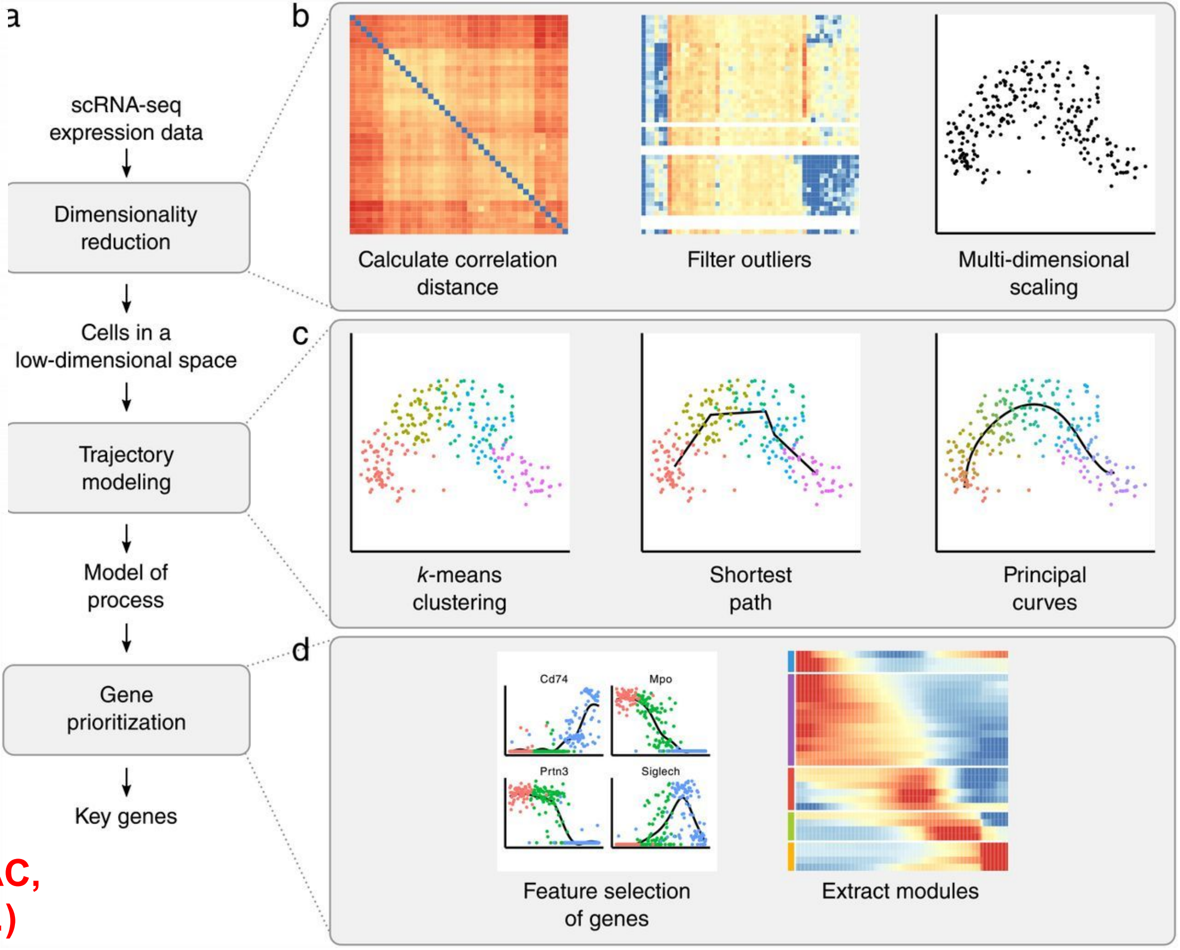
Differential expression analysis : methods



Soneson & Robinson 2018

Cell trajectory : methods





SCORPIUS, bioRxiv (2016)

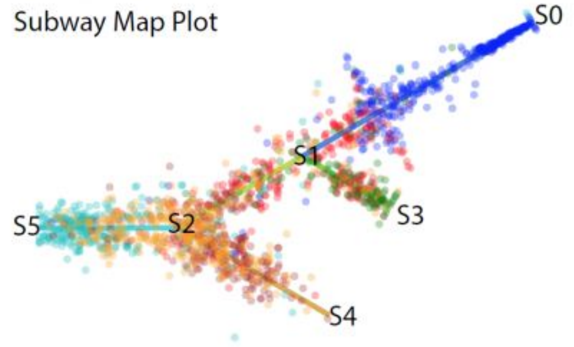
Most adopted tools :

- Monocle 3
- PAGA
- STREAM
- Scorpius
- Slingshot

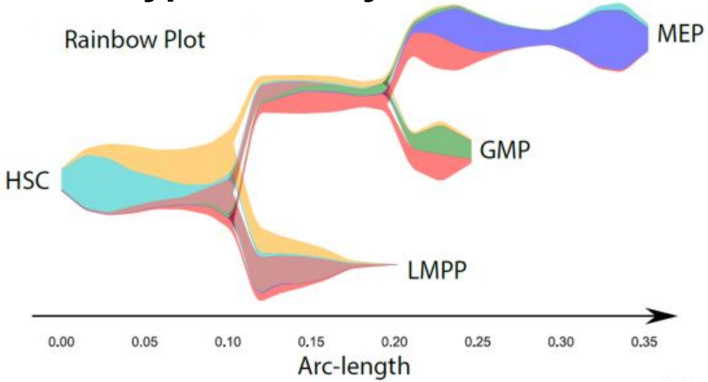
Not limited to scRNAseq ! (ATAC, CITE, multiomics, imagery-based …)

Cell trajectory : visualization

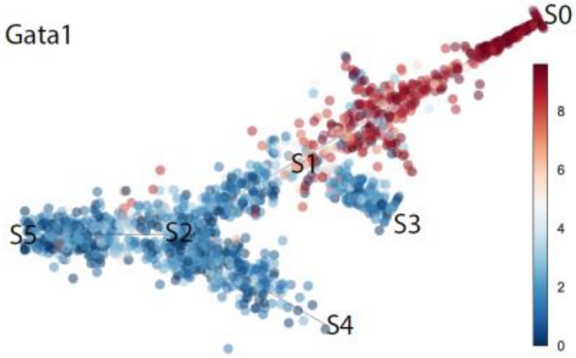
Cell distance to path + cell types



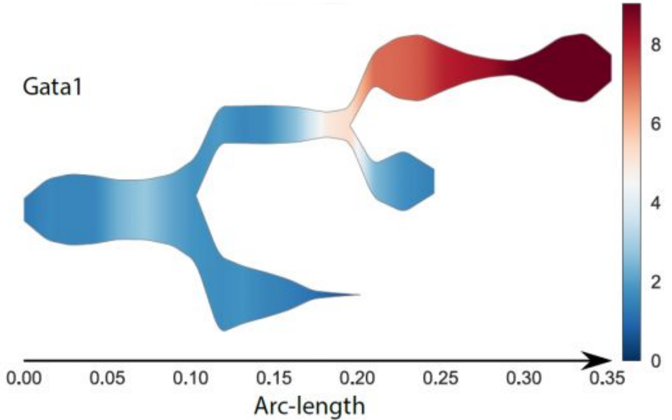
Cell types density



Cell distance to path + gene expression

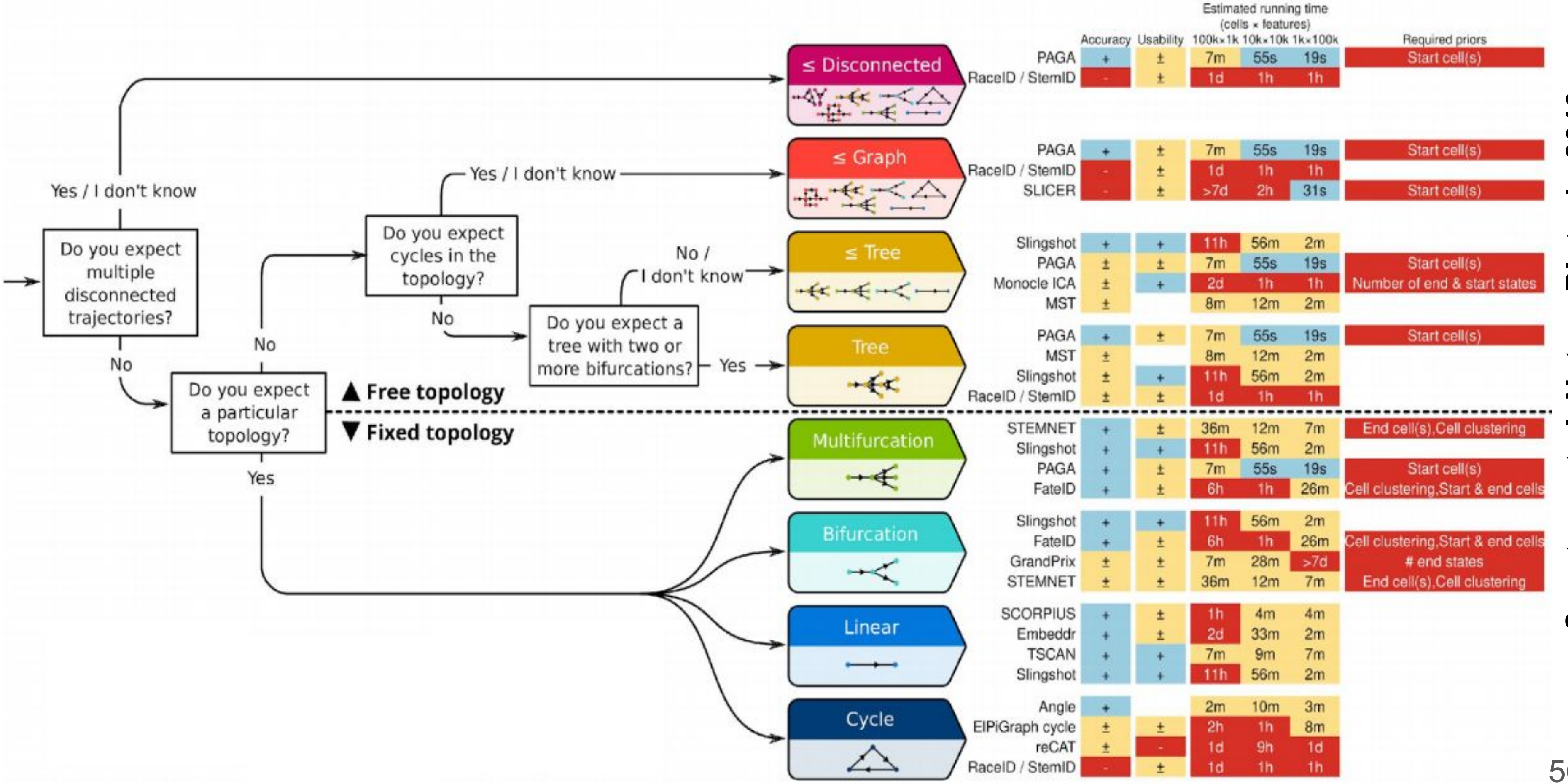


Cells density + gene expression



STREAM (Pinello Lab)

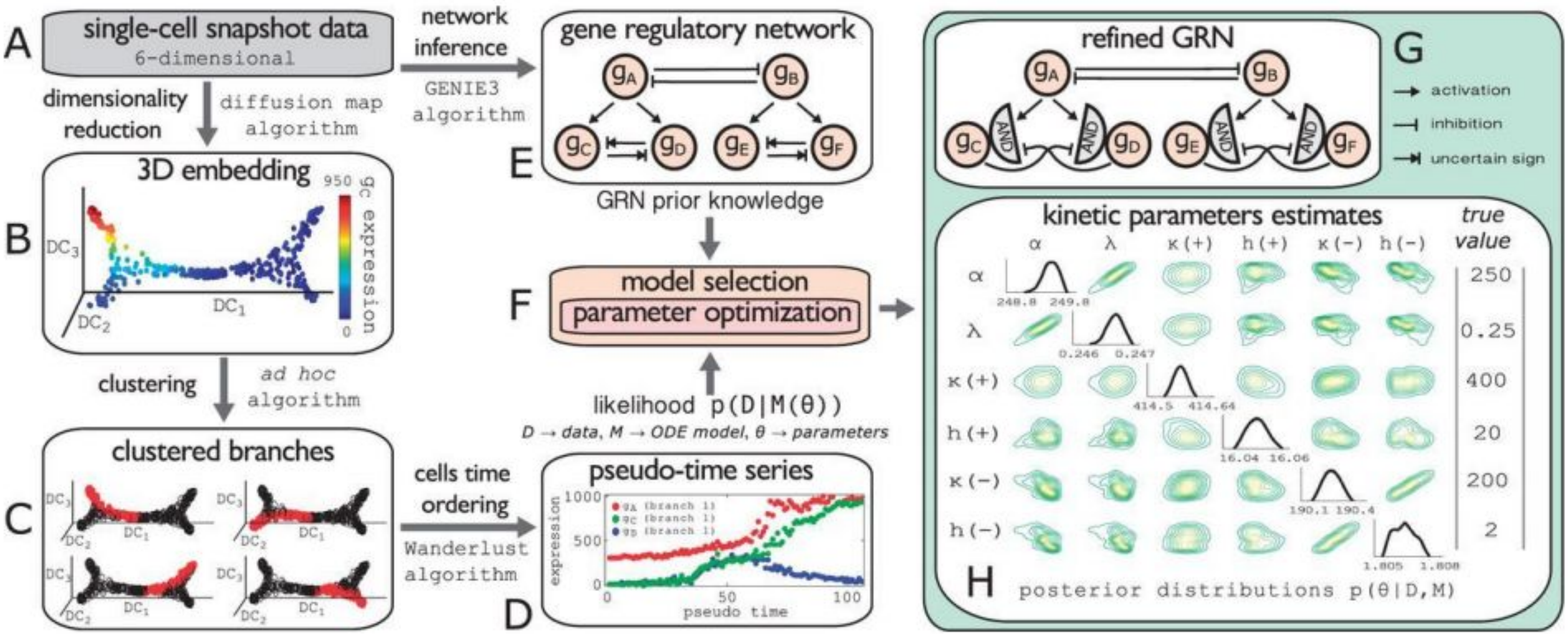
Cell trajectory : Contexts



Saoutens et al, Nature Biotech, 2019

Network inference

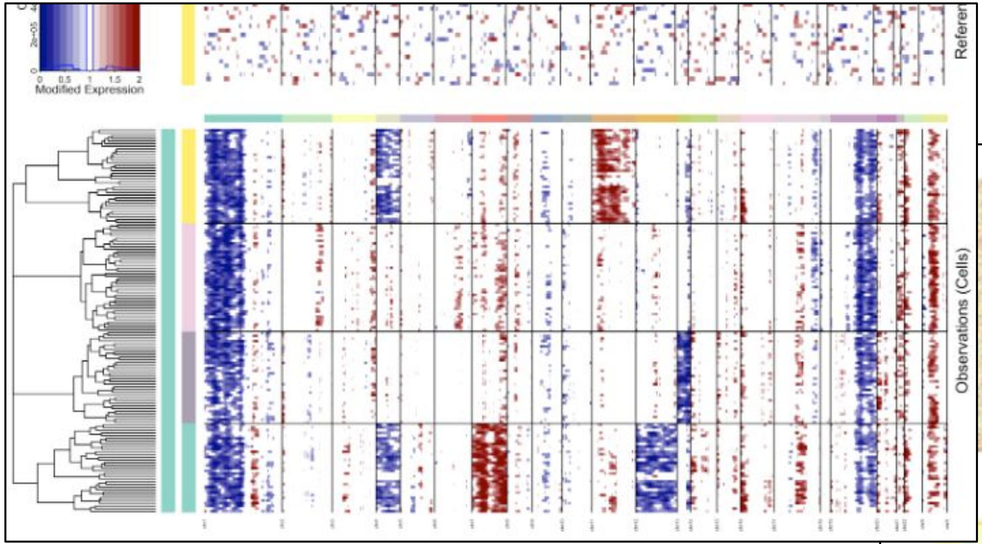
Using cell ordering from trajectory analysis + co-occurring / correlated genes



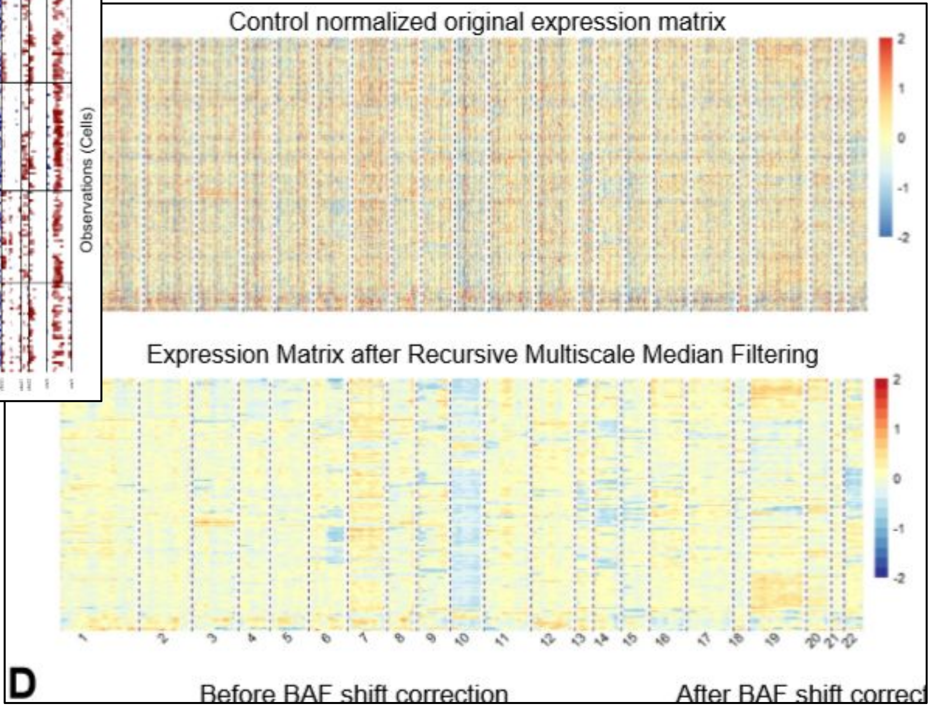
Ocone et al. Bioinformatics (2015)

Copy number estimation from scRNAseq

InferCNV (Broad Institute)



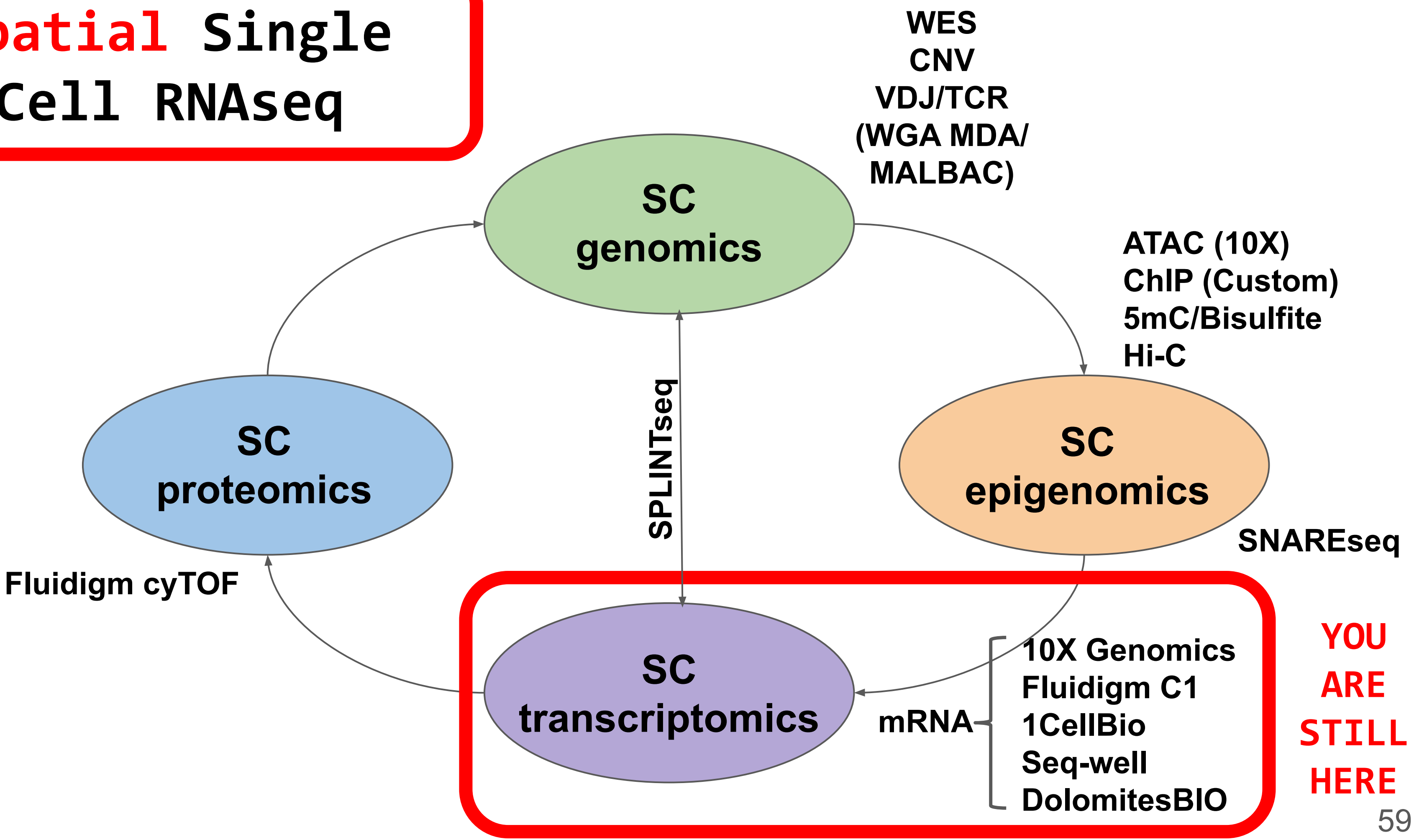
CaSpER (Armanci et al, BioRxiv 2019)

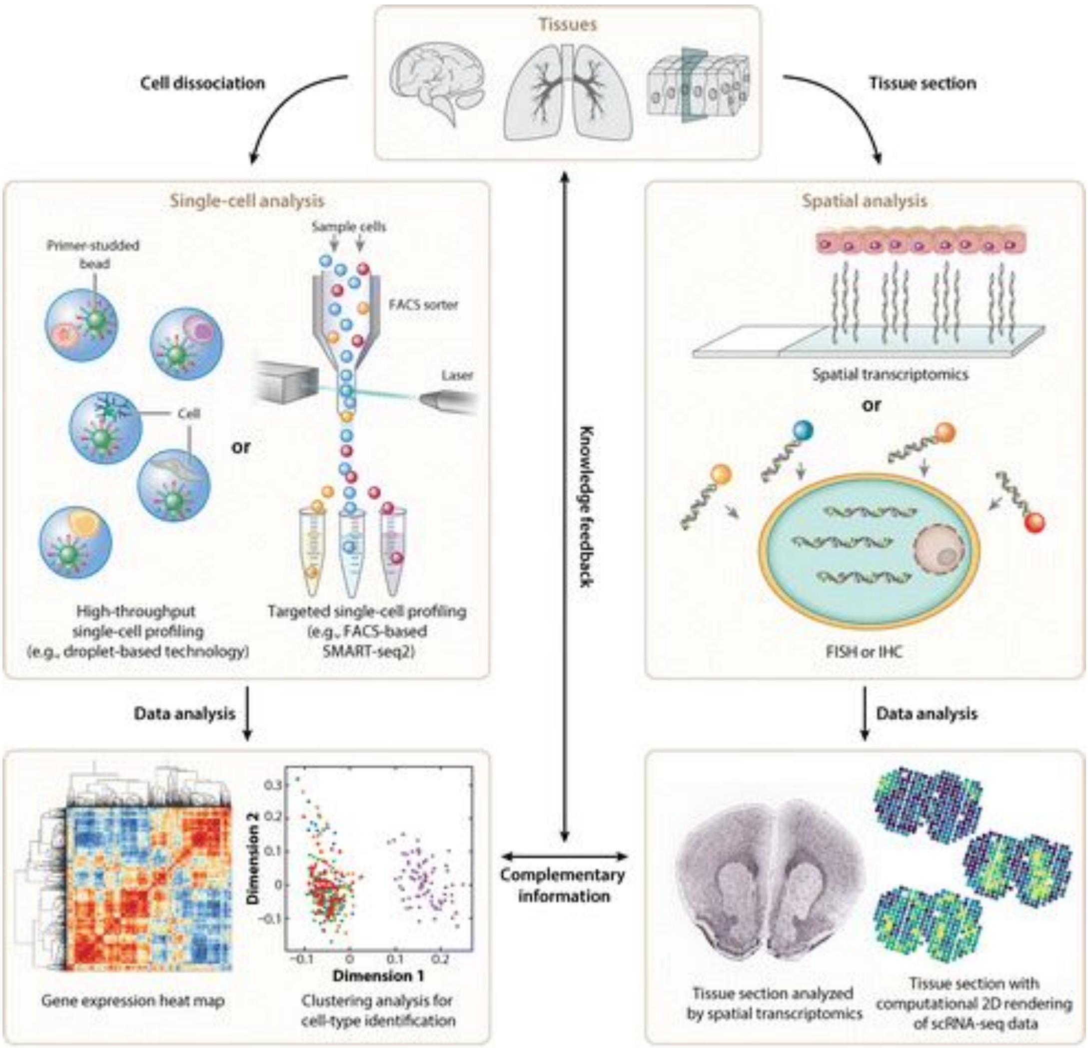


WARNING :

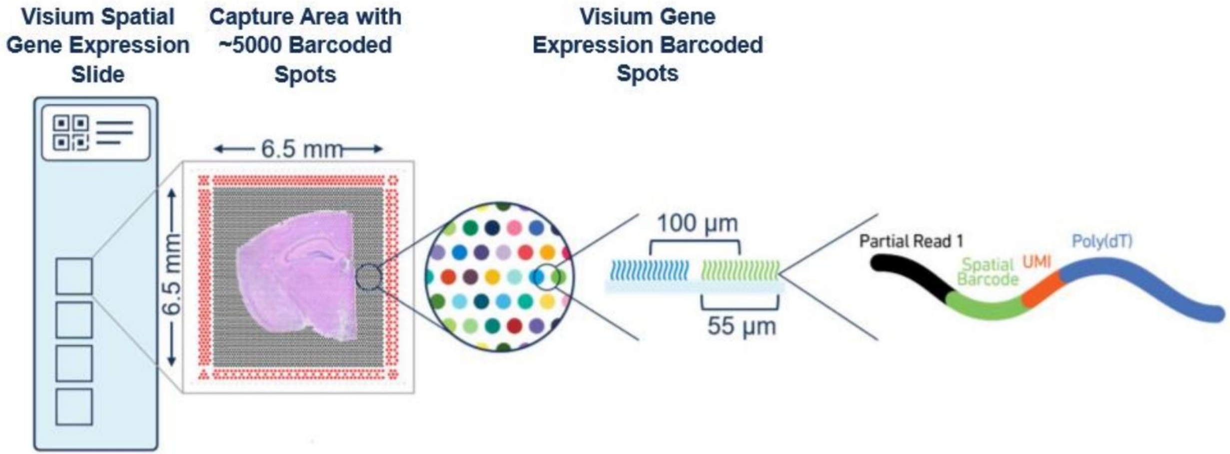
- Coarse grain (> 10 Mb)
- Requires > 75,000 reads / cell

Spatial Single Cell RNAseq



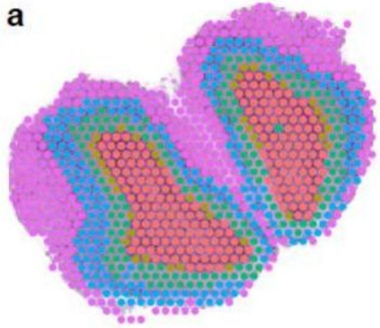

Tissues
Cell dissociation
Tissue section
Single-cell analysis
Primer-studded bead
Sample cells
FACS sorter
Laser
Cell
or
High-throughput single-cell profiling (e.g., droplet-based technology)
Targeted single-cell profiling (e.g., FACS-based SMART-seq2)
Spatial analysis
Spatial transcriptomics
or
FISH or IHC
Knowledge feedback
Data analysis
Data analysis
Dimension 2
Dimension 1
Gene expression heat map
Clustering analysis for cell-type identification
Complementary information
Tissue section analyzed by spatial transcriptomics
Tissue section with computational 2D rendering of scRNA-seq data

10x Genomics Visium

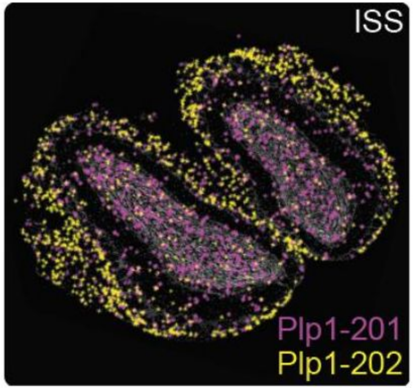


IHC

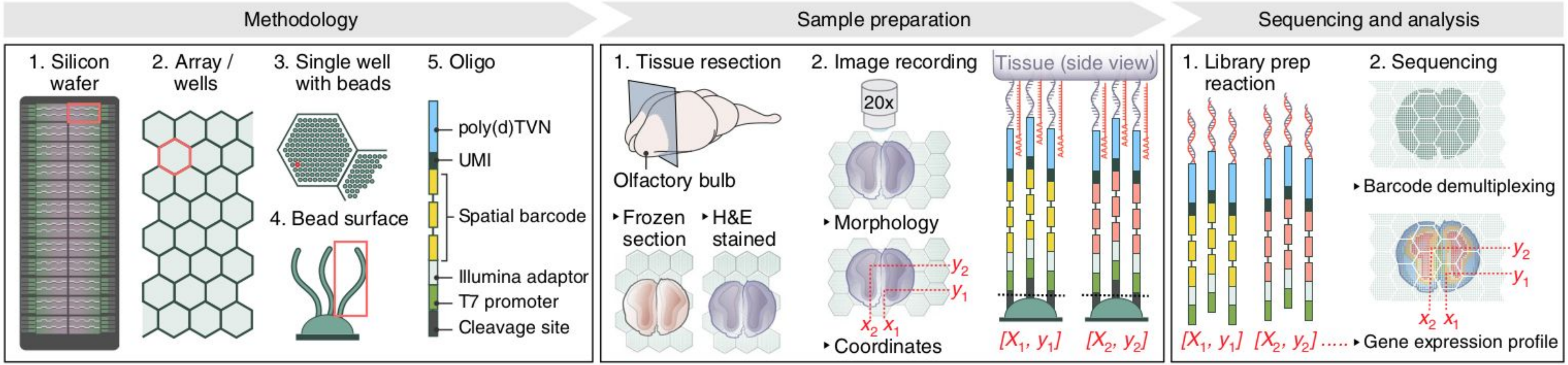
Clusters



RNA FISH

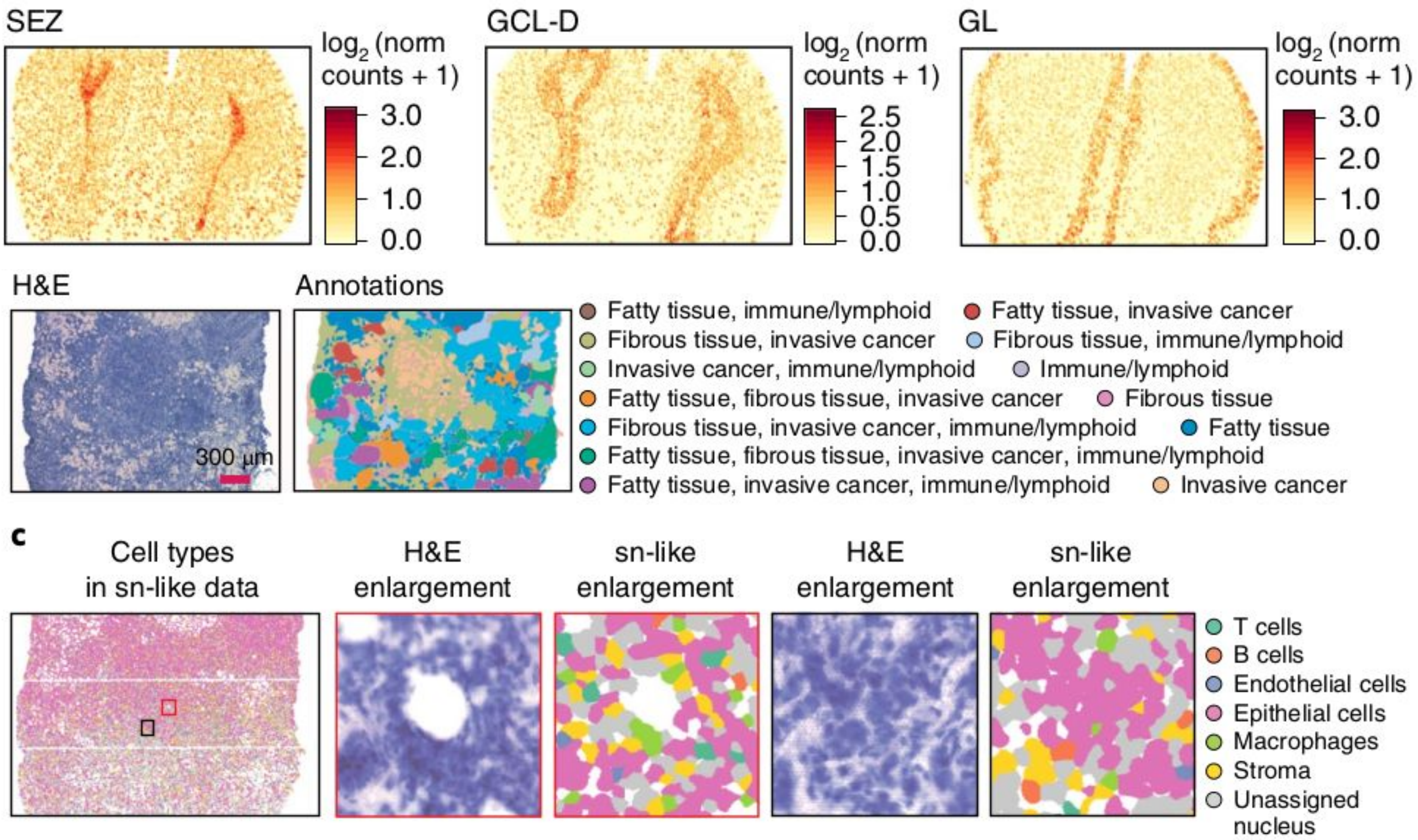


Illumina “HD Spatial Transcriptomics”

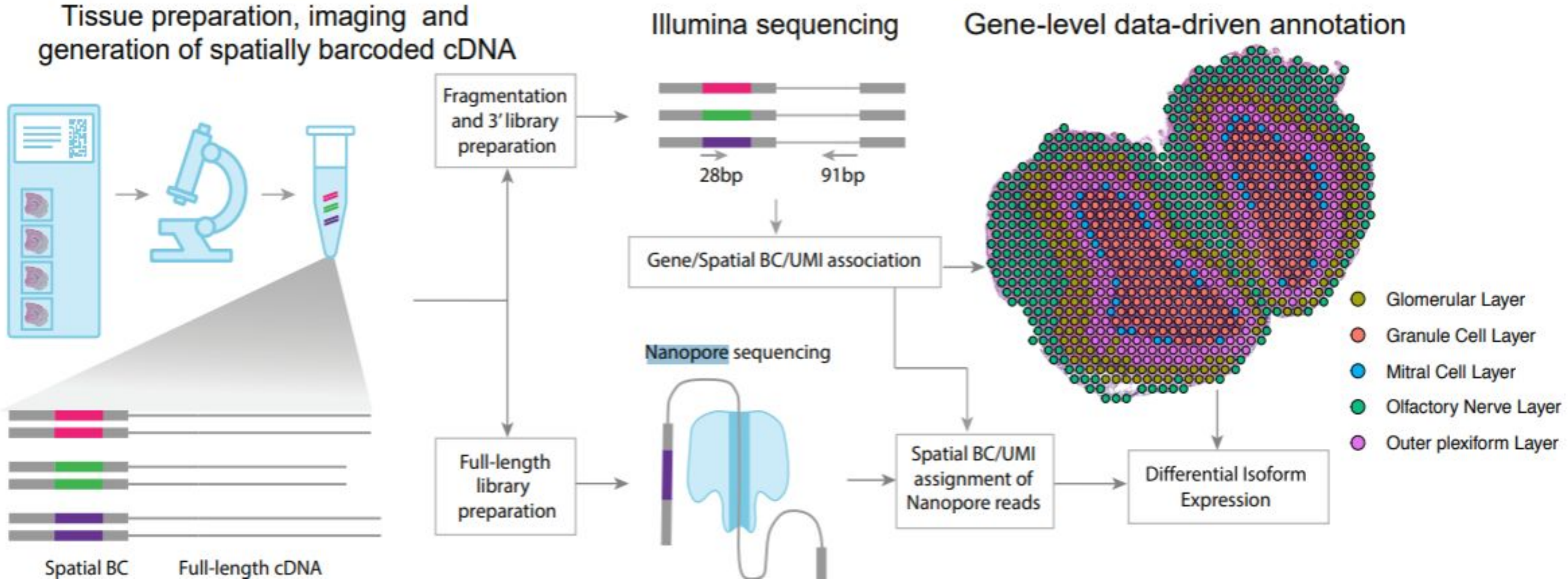


- **2,893,865 individual barcoded beads**
- **1,4 M wells**
- **Well diameter ~ 2 µm**
  - **<< median cell diameter (20 µm)**
  - **~ 1,400 x higher resolution than “standard” ST**
  - **~ 25 x compared to SLIDE-seq**
- **Array reading time ~ 3 H**
- **Challenging analysis strategy (low capture rate) …**
- **Commercially available in 2020**

Illumina HDST



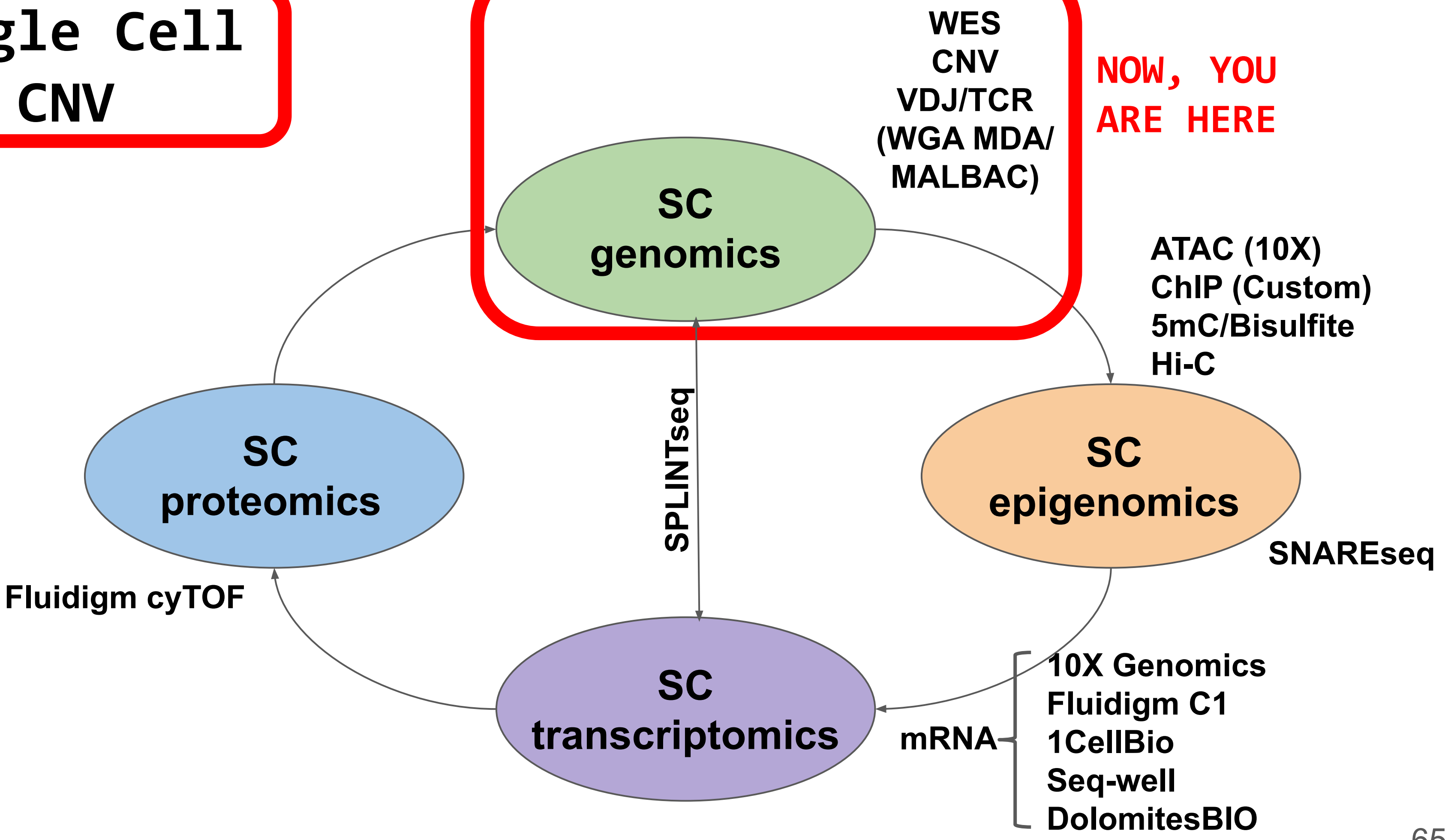
Spatial long reads



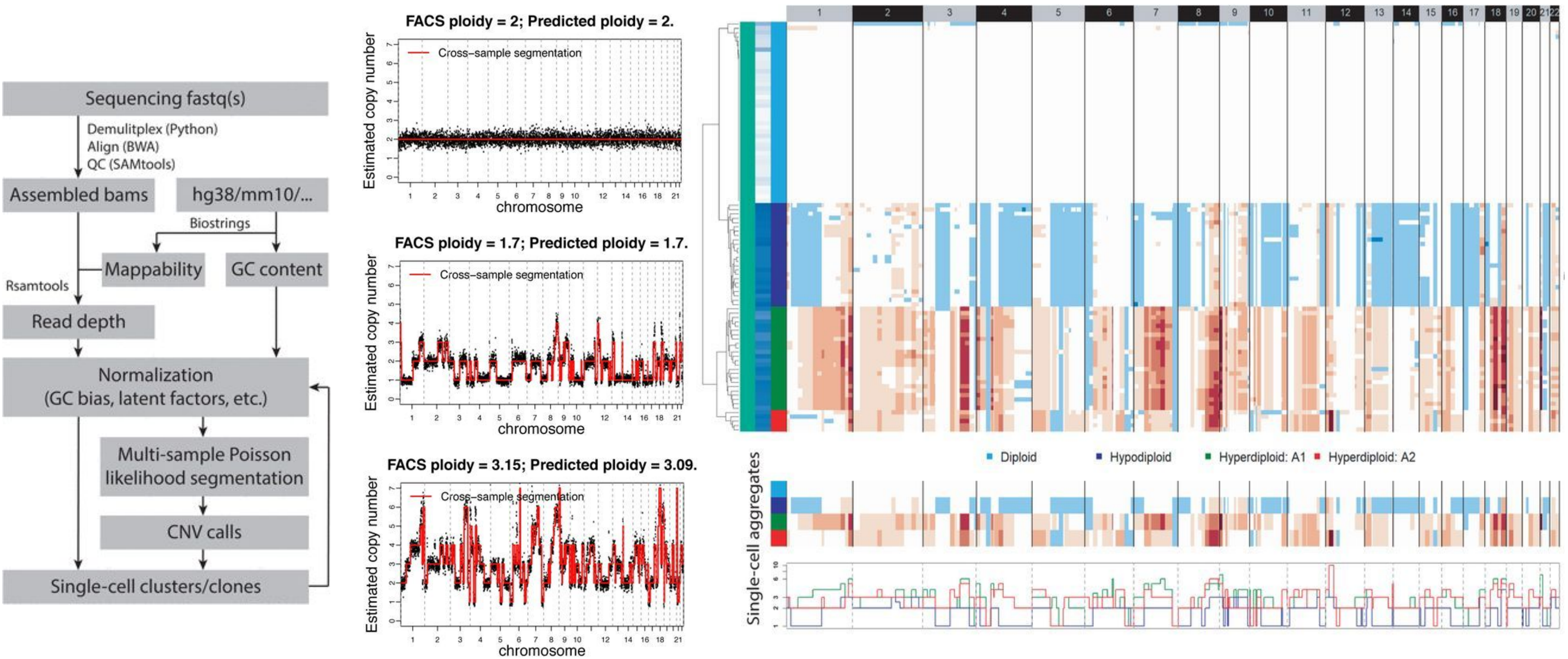
Lebrigand et al, BiorXiv, 2020/08

Single Cell CNV

NOW, YOU ARE HERE



scCNV results (SCOPE)



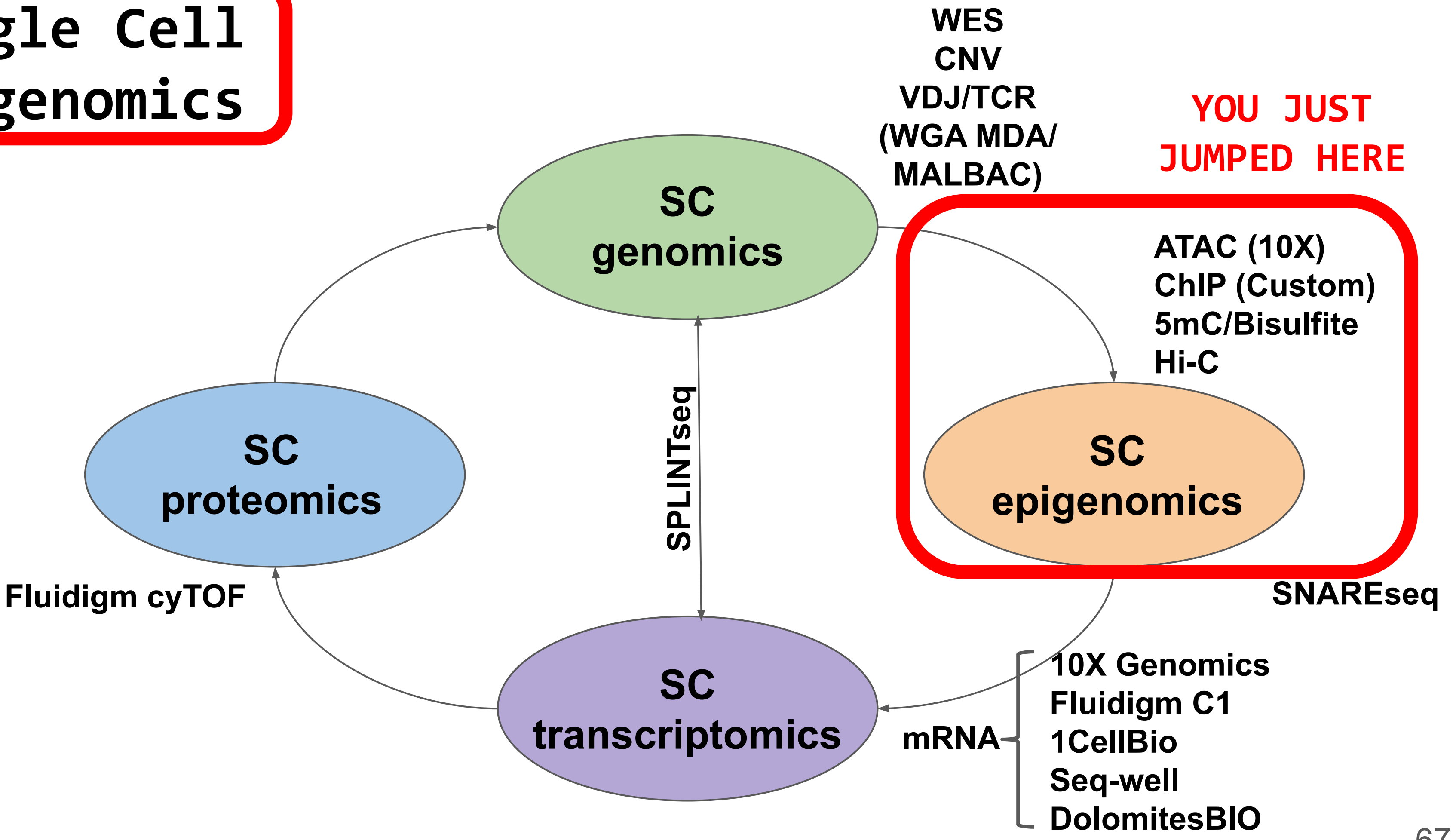
SCOPE Rujn Wang et al., bioRxiv, 2019/03

WARNING :

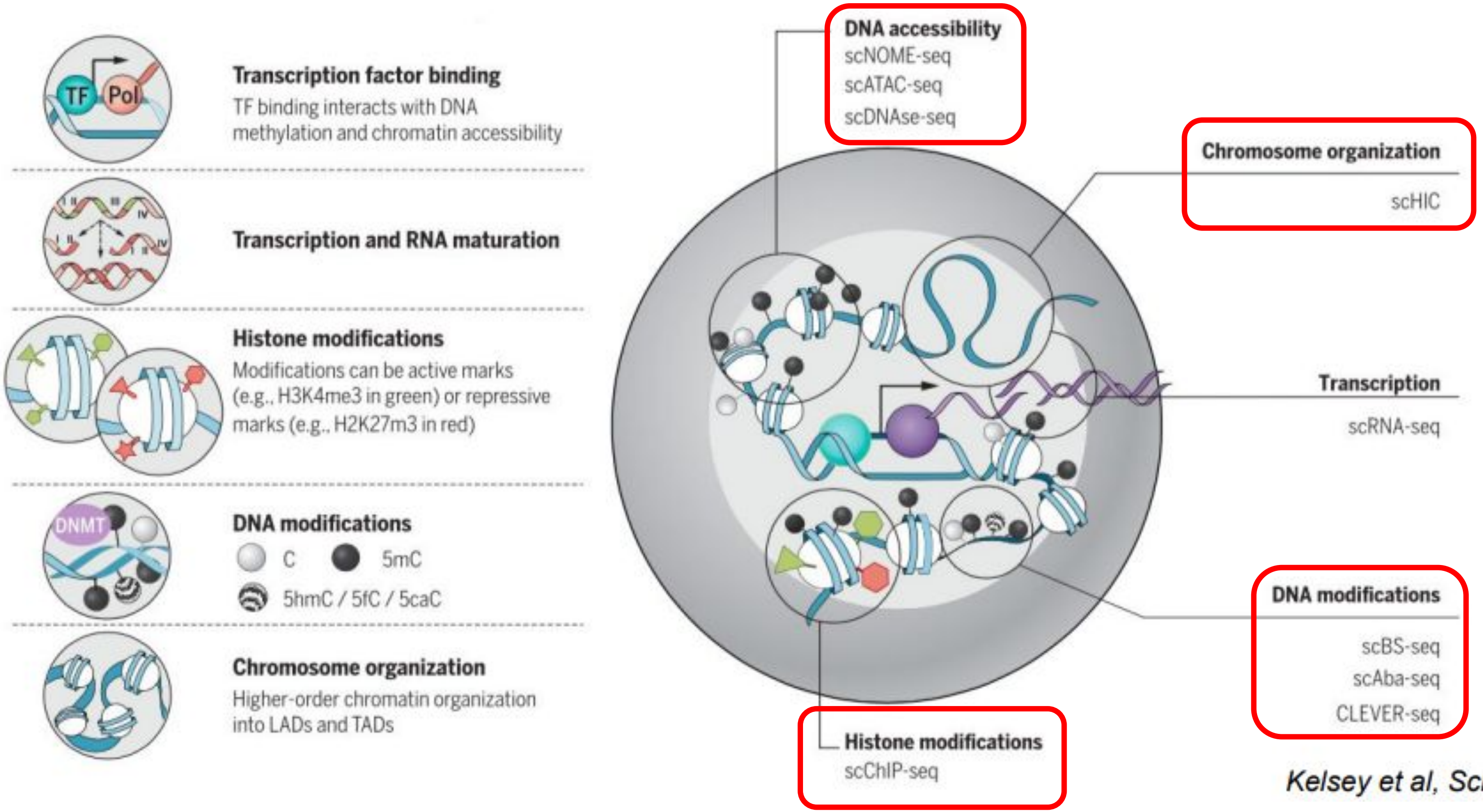
- Limited resolution : > 2 Mb (binning)
- Requires > 750,000 reads / cell

ALSO available : ***SCYN***

Single Cell Epigenomics

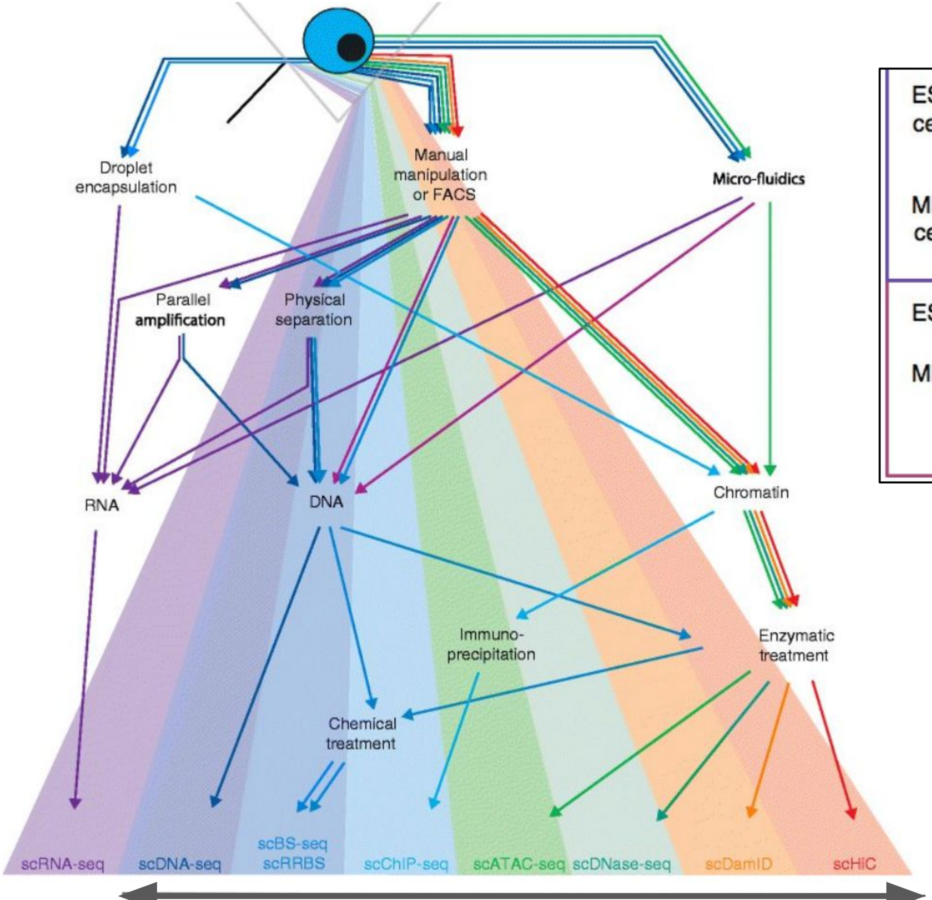


Overview of scEpigenomics techniques

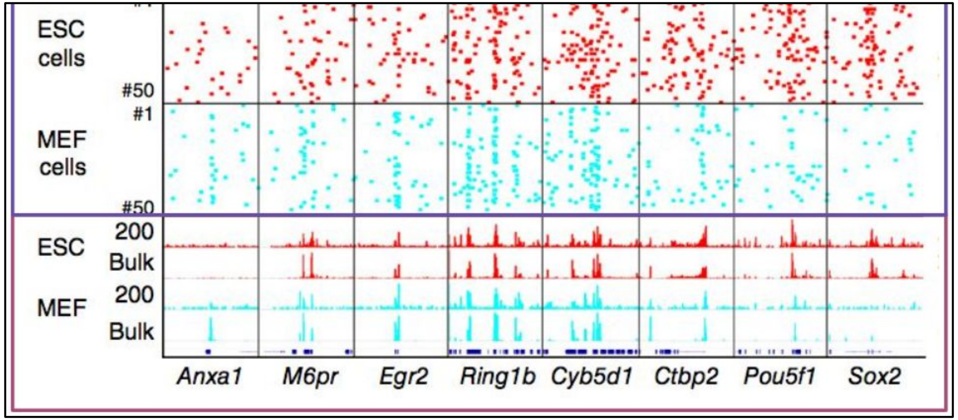


Kelsey et al, Science, 2017

Overview of scEpigenomics techniques



scChIPseq



- scChIP : improvements in 2019
- scMeth : low coverage, low sensitivty (<20% CpG read)
- scHi-C : stable protocol & analysis still needed
- scATAC : most popular technology, numerous tools available

Single cell (RNAseq) resources

(some)

Tabula Muris

ARTICLE

https://doi.org/10.1038/s41586-018-0590-4

Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*

The Tabula Muris Consortium*

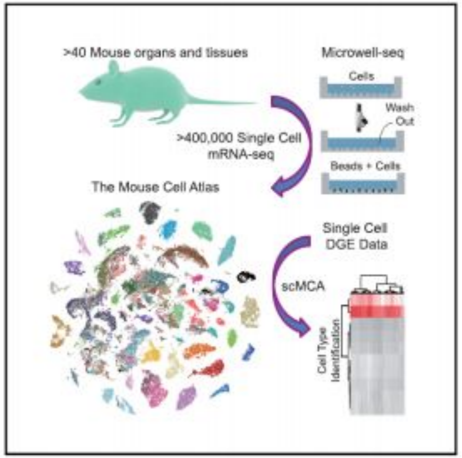
- ~100k cells
- 20 organs
- 2 techniques :
  - Droplet 3’, short reads
  - FACS, long reads

MCA browser
http://bis.zju.edu.cn/MCA/

Cell — Resource

Mapping the Mouse Cell Atlas by Microwell-Seq

Graphical Abstract

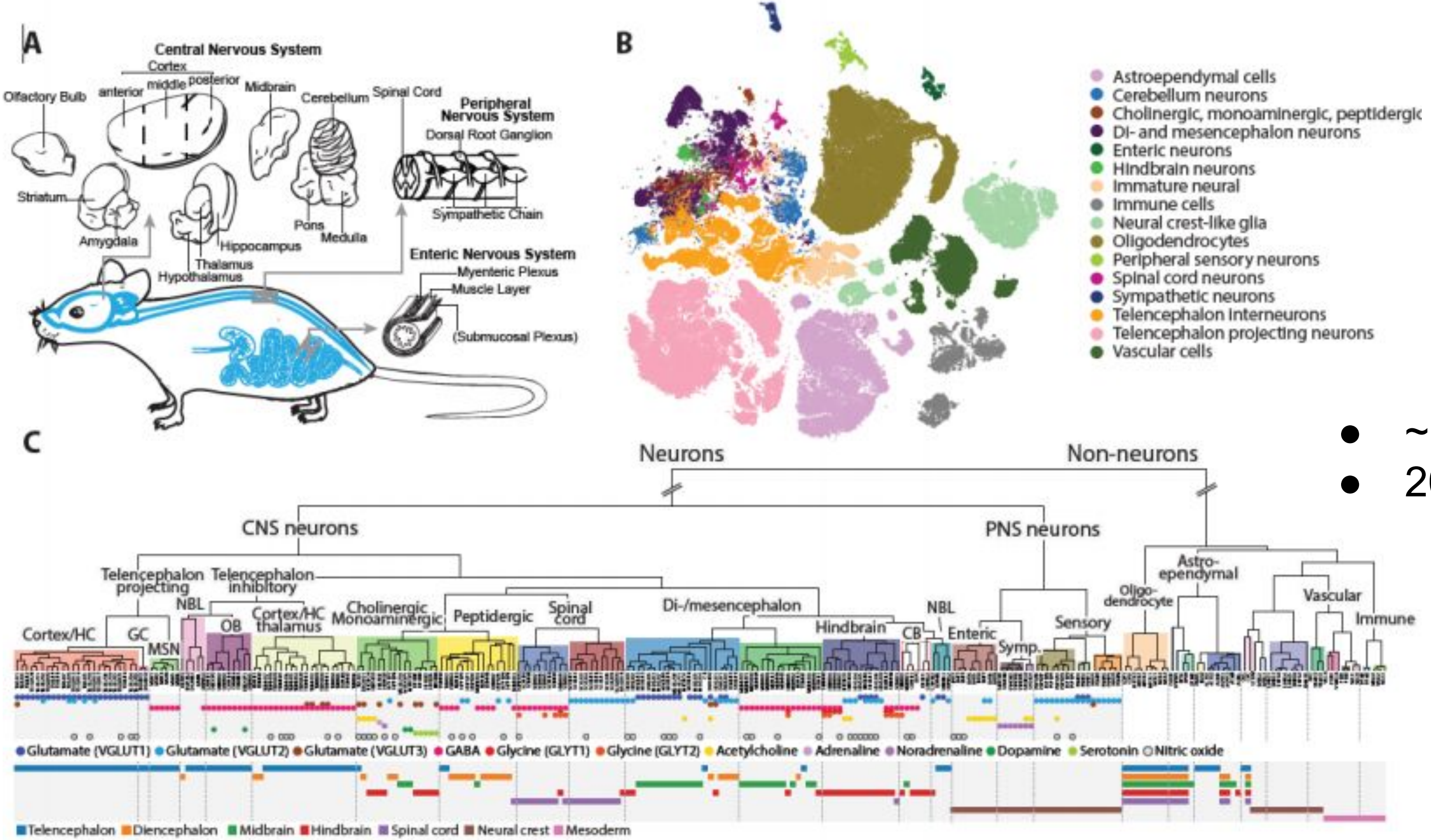


Authors
Xiaoping Han, Renying Wang, Yincong Zhou, ..., Guo-Cheng Yuan, Ming Chen, Guoji Guo

Correspondence
xhan@zju.edu.cn (X.H.), ggj@zju.edu.cn (G.G.)

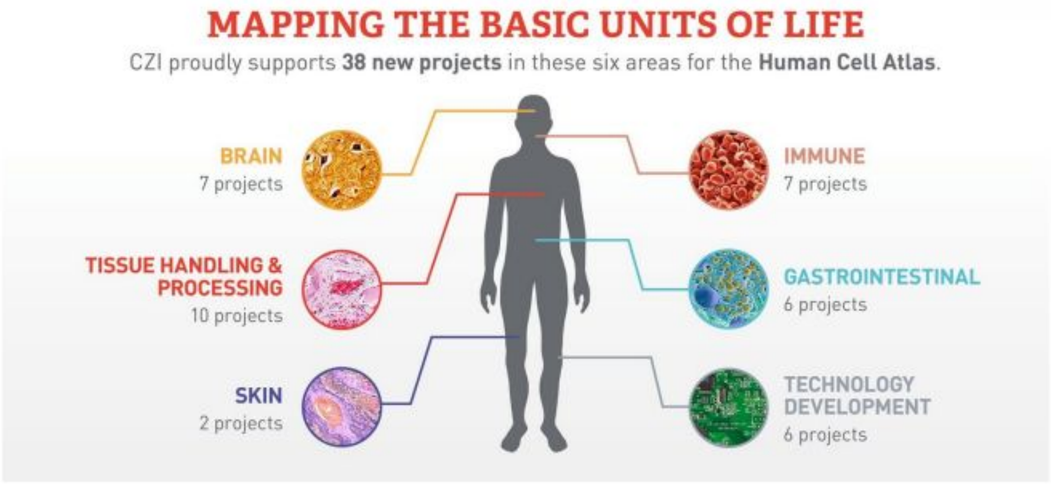
In Brief
Development of Microwell-seq allows construction of a mouse cell atlas at the single-cell level with a high-throughput and low-cost platform.

The Mouse Brain Atlas (mousebrain.org)



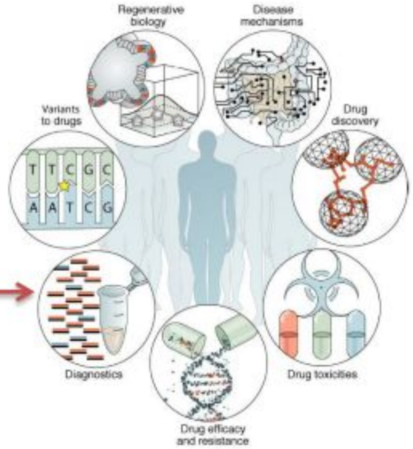
- ~160k cells
- 20 subtypes

The Human Cell Atlas (humancellatlas.org)



HUMAN CELL ATLAS

- Every cell type in the body
- First: define how to proceed
  - Best experimental practice / organ
  - Best bioinformatics methods
- Data will be made available to all



PanglaoDB

PanglaoDB | Home | Search | Datasets | Tools | Papers | FAQ/Help | About

PanglaoDB is a database for the scientific community interested in exploration of single cell RNA sequencing experiments from mouse and human. We collect and integrate data from multiple studies and present them through a unified framework.

Usage examples

- Run a gene search for SOX2, PECAM1 or ACE2
- Browse the full list of samples
- Explore the list of cell type markers for Schwann cells
- Browse cell types of the mouse retina
- Look at the expression of *CRX* in photoreceptor cells

Database statistics

	Mus musculus	***Homo sapiens***
Samples	1063	305
Tissues	184	74
Cells	4,459,768	1,126,580
Clusters	8,651	1,748

Dataset of the day

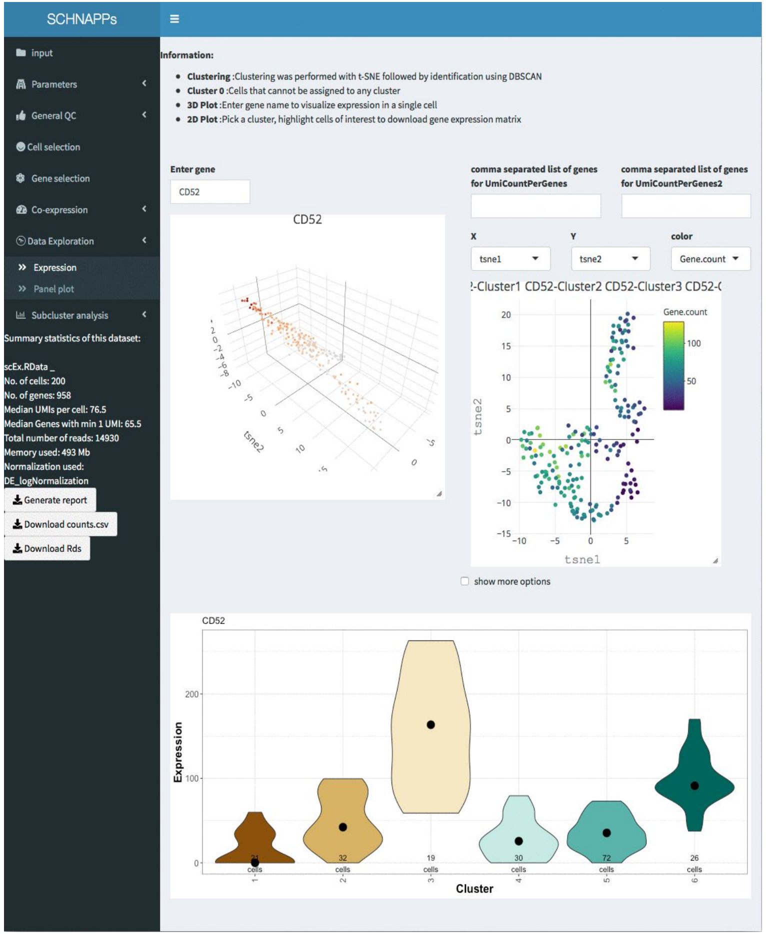
Take a closer look at the cellular composition of Heart, using a dataset which consists of 871 cells. Clustering of this dataset resulted in 7 cell clusters, containing among others, Smooth

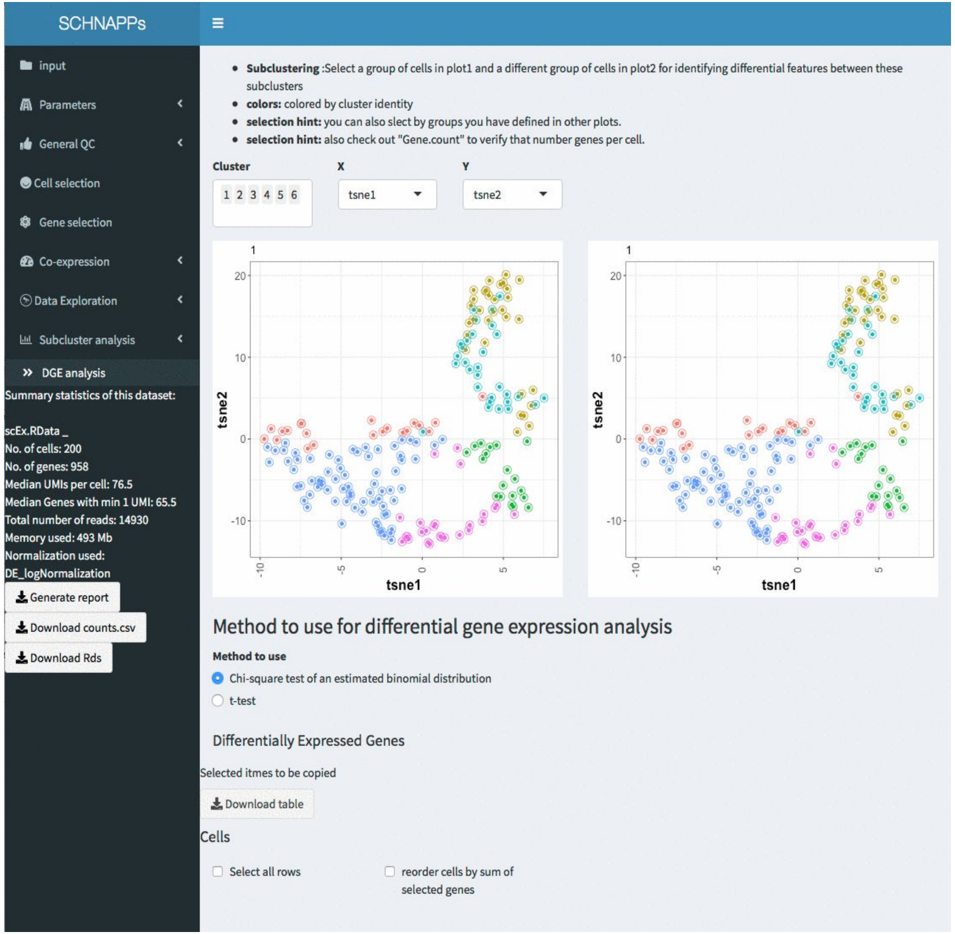
https://panglaodb.se/

WYSIWYG analysis frameworks

(mainly for scRNAseq / scATACseq)

SCHNAPPS : A R-shiny app for biologists



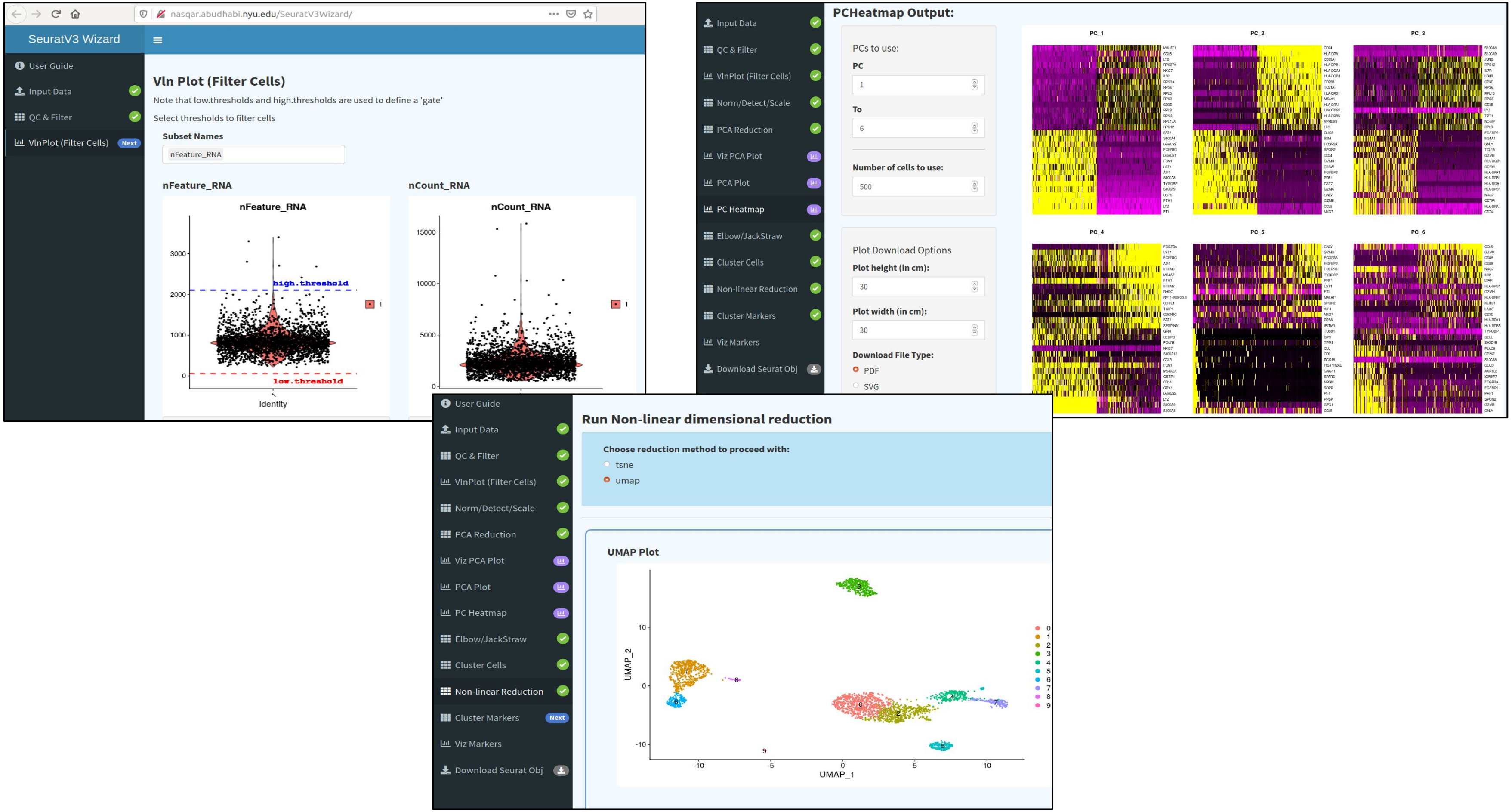


By Bernd Jagla (Pasteur Paris)

https://c3bi-pasteur-fr.github.io/UTechSCB-SCHNAPPs

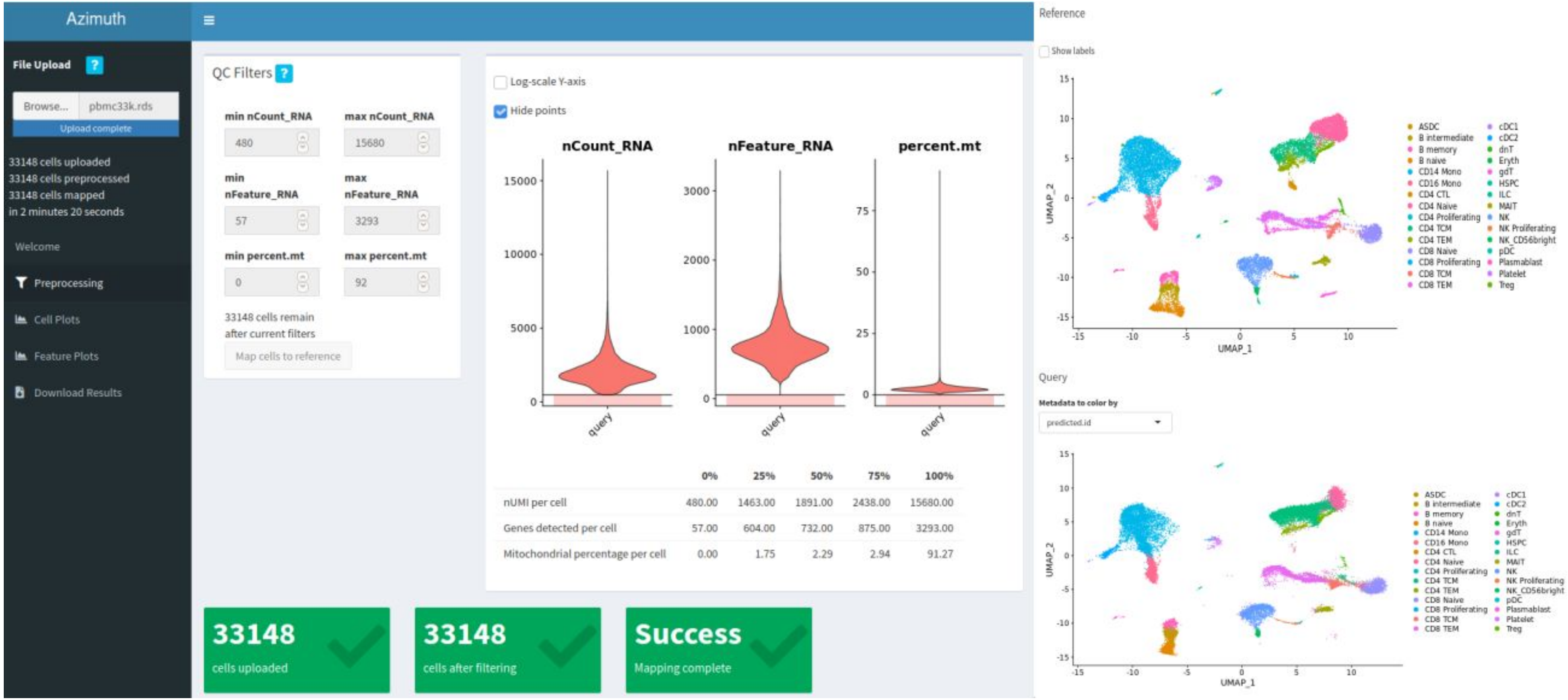
https://github.com/C3BI-pasteur-fr/UTechSCB-SCHNAPPs

SeuratV3Wizard



http://nasqar.abudhabi.nyu.edu/SeuratV3Wizard/

Azimuth (from Satija Lab)

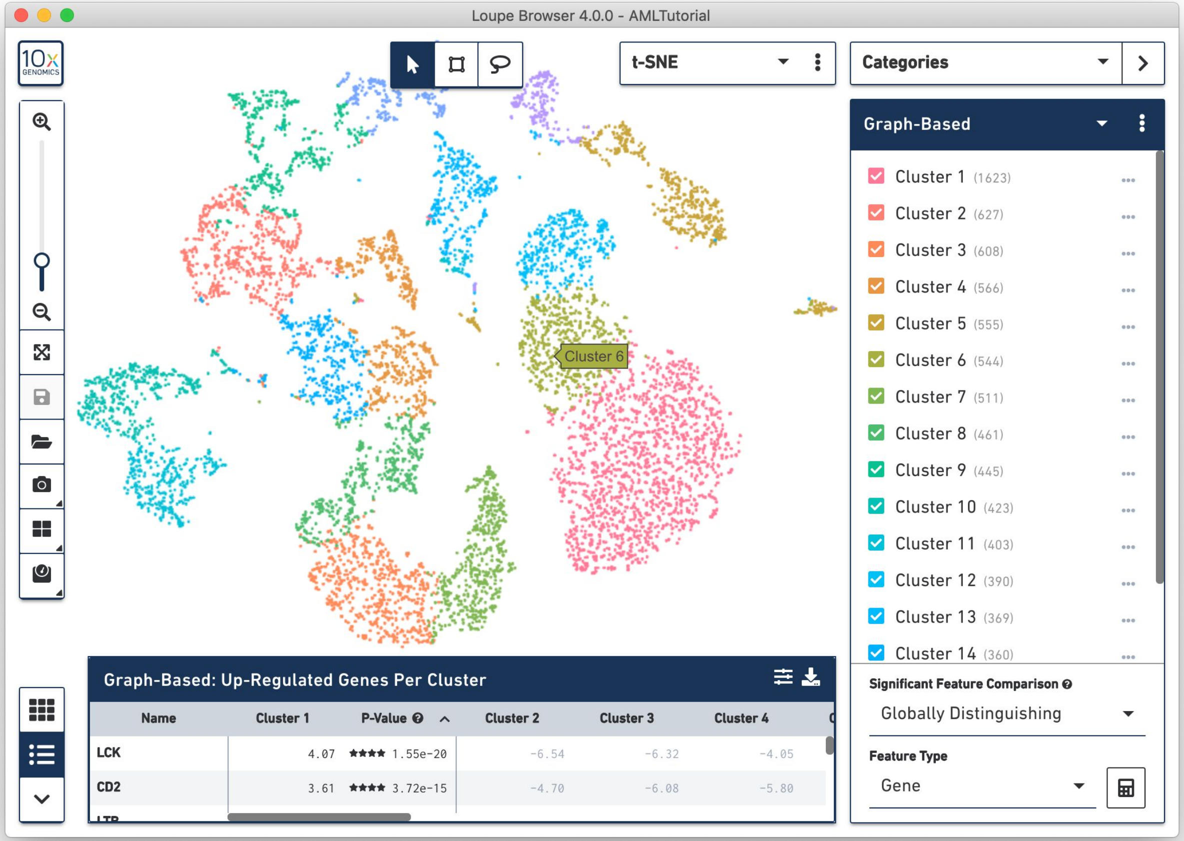


https://satijalab.org/azimuth/#General

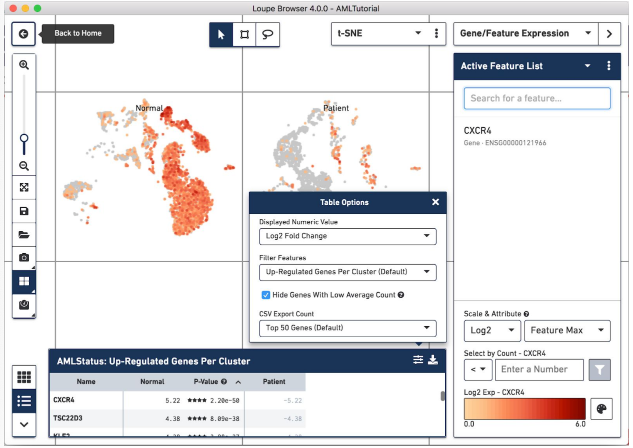
Visualization tools

(mainly for scRNAseq)

10x Genomics Loupe Browser



- Compatible with output from 10x Cell Ranger (“cloupe” files)
- Linux / OSX
- Supports Visium (Spatial)

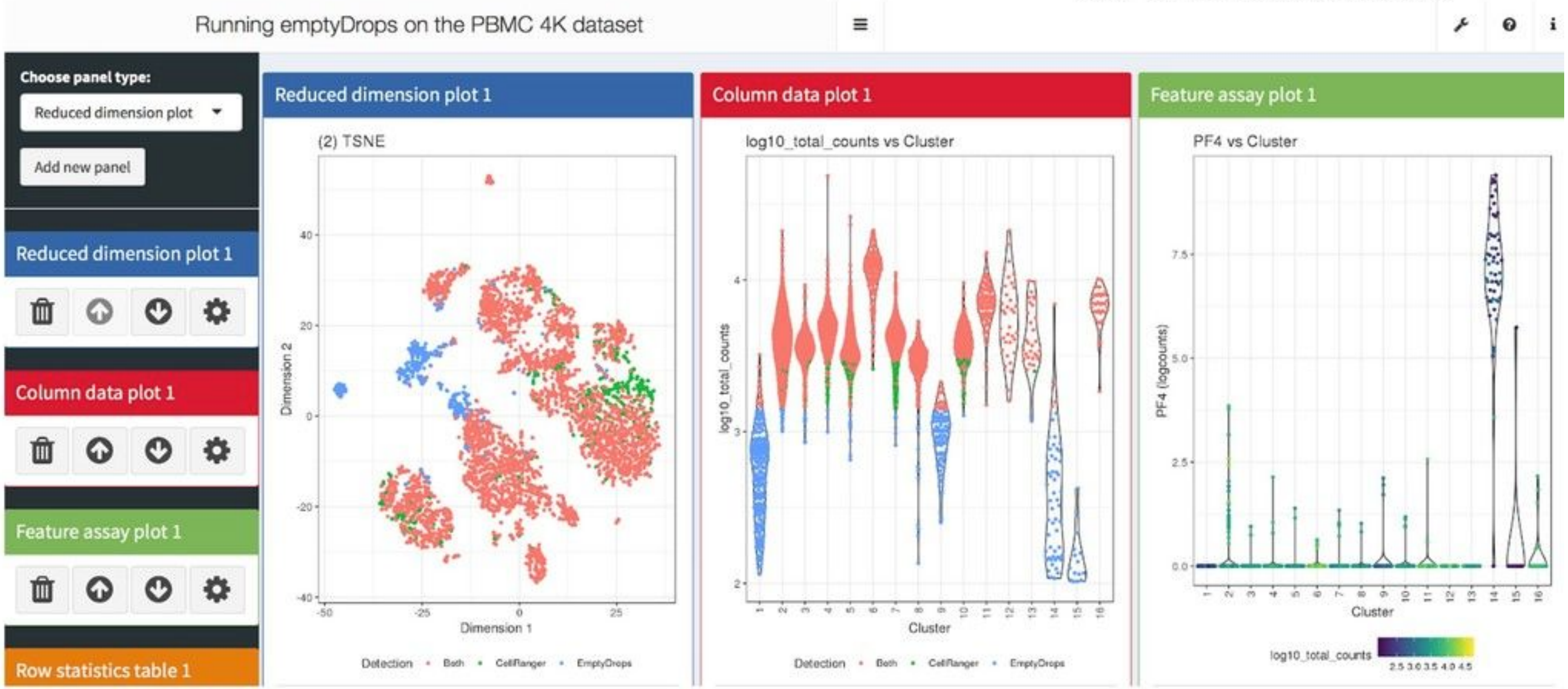


https://support.10xgenomics.com/single-cell-gene-expression/software/visualization/latest/what-is-loupe-cell-browser

iSee

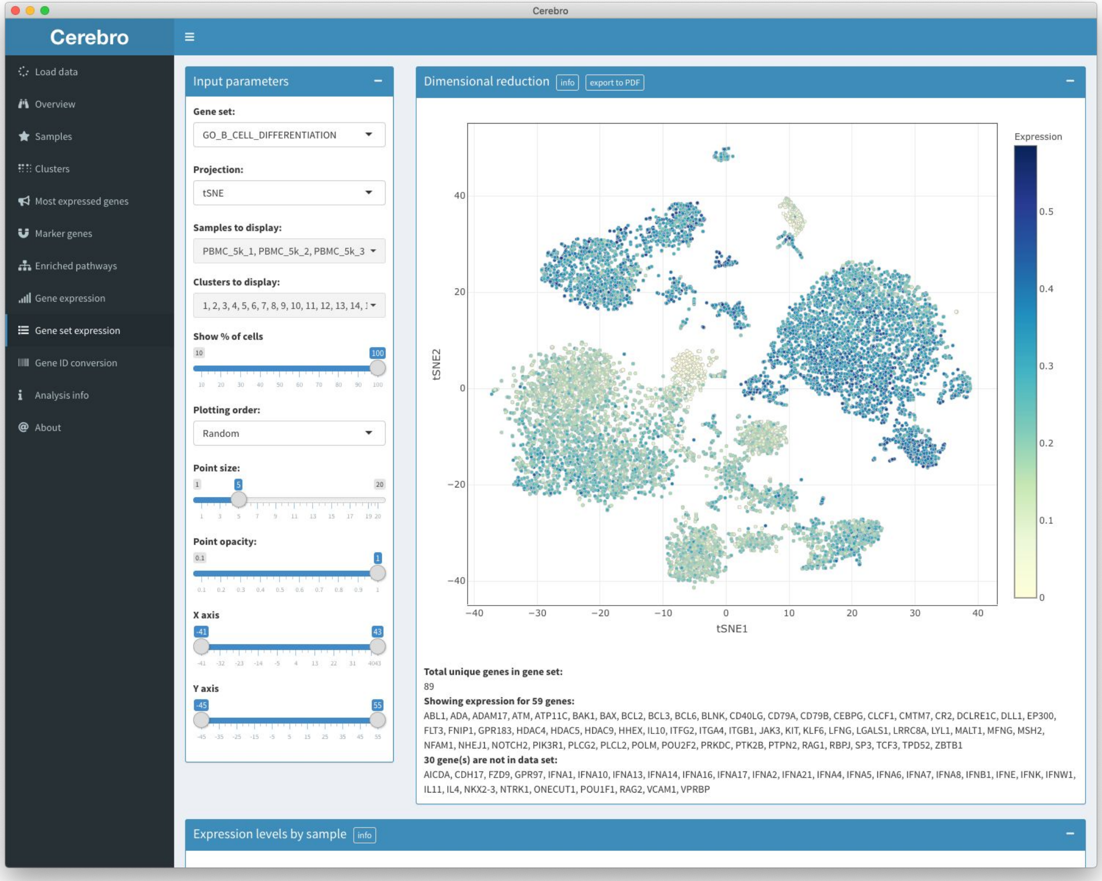
Interactive Data Visualization (iSEE)

Creators: Federico Marini, Aaron Lun, Charlotte Soneson, and Kevin Rue-Albrecht



https://marionilab.cruk.cam.ac.uk/iSEE_pbmc4k/

CerebroApp

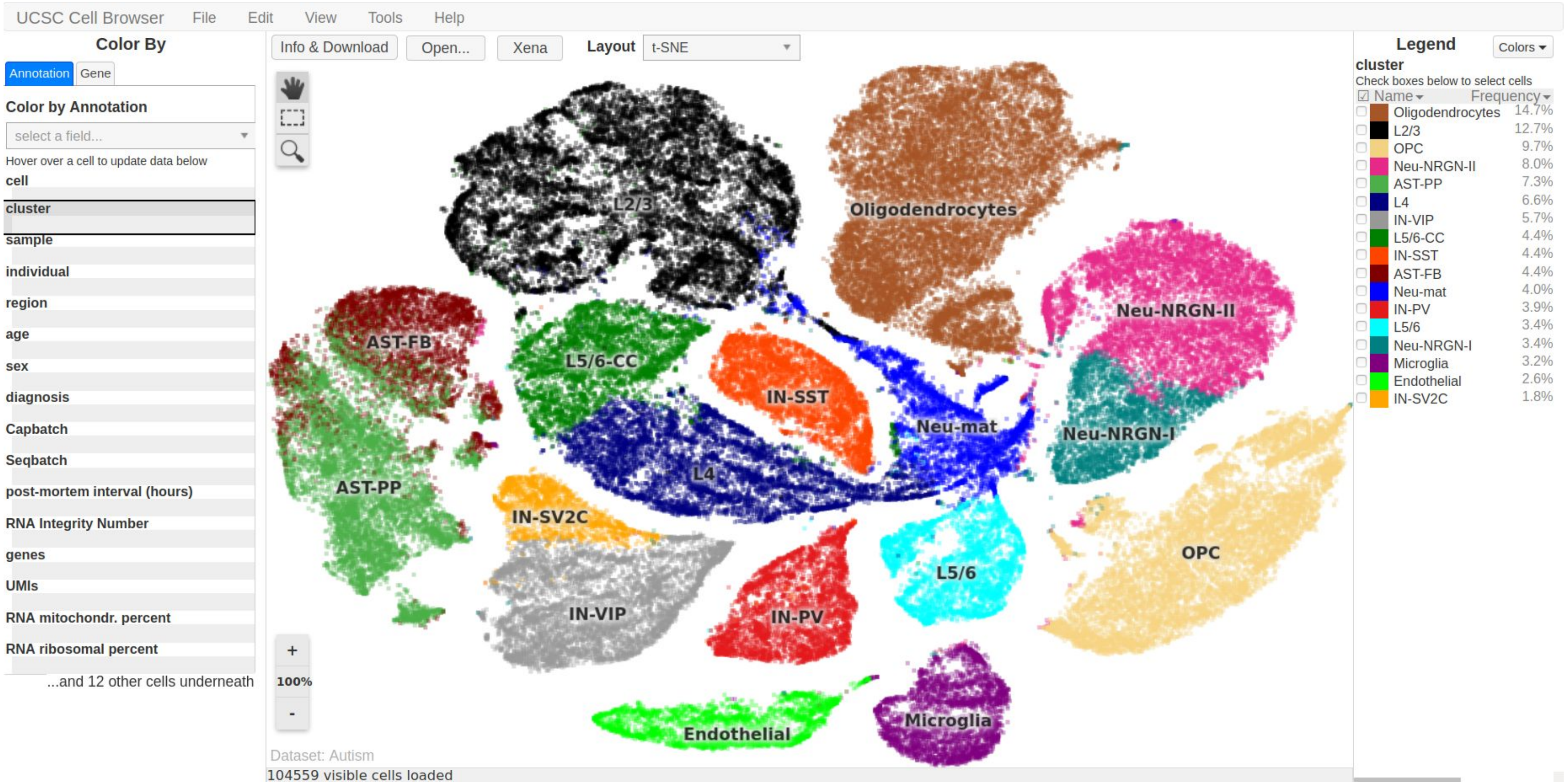


- ShinyApp (web GUI over some R)
- Binary format (CRB), converted from SeuratObject / SCE
- From QC to trajectory

(my favorite one)

https://romanhaa.github.io/projects/cerebro/

UCSC Cell Browser



https://github.com/maximilianh/cellBrowser (demo : https://cells.ucsc.edu)

BREAKING NEWS !

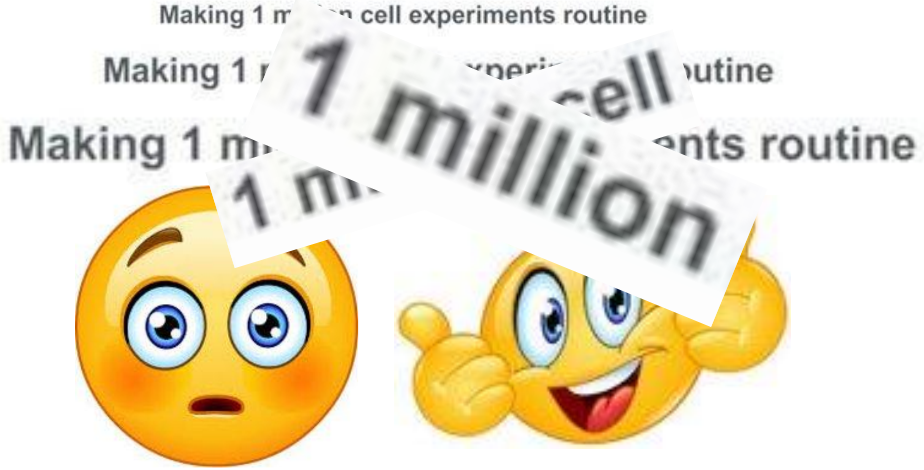
Announced by 10x on *twitter* in January

New High Throughput Instrument

Chromium X

Coming 2H 2021

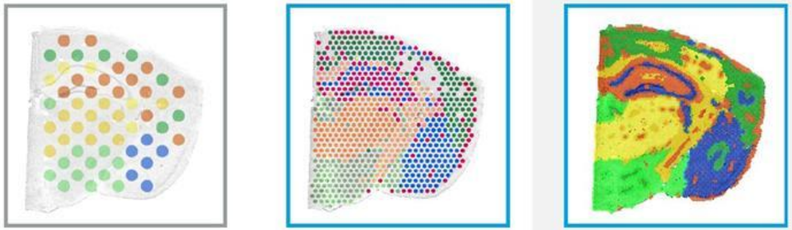
Making 1 million cell experiments routine



Visium at Single Cell Resolution

Visium HD

Coming 1H 2022



First Generation

Visium
Resolution 4x

Visium HD
Resolution 1,500x

Unlocking More Samples for Spatial

Visium CytAssist

Coming 1H 2022

Accessing archived and pre-mounted samples

Acknowledgements

Marc Deloger
Morgane Thomas-Chollier
Agnès Paquet
Marine Aglave
Antonio Rausell
Wouter Saelens
Nathalie Gaspar

SINGle-cellING in the RAINaseq (1952)

A joke © Jacques van Helden

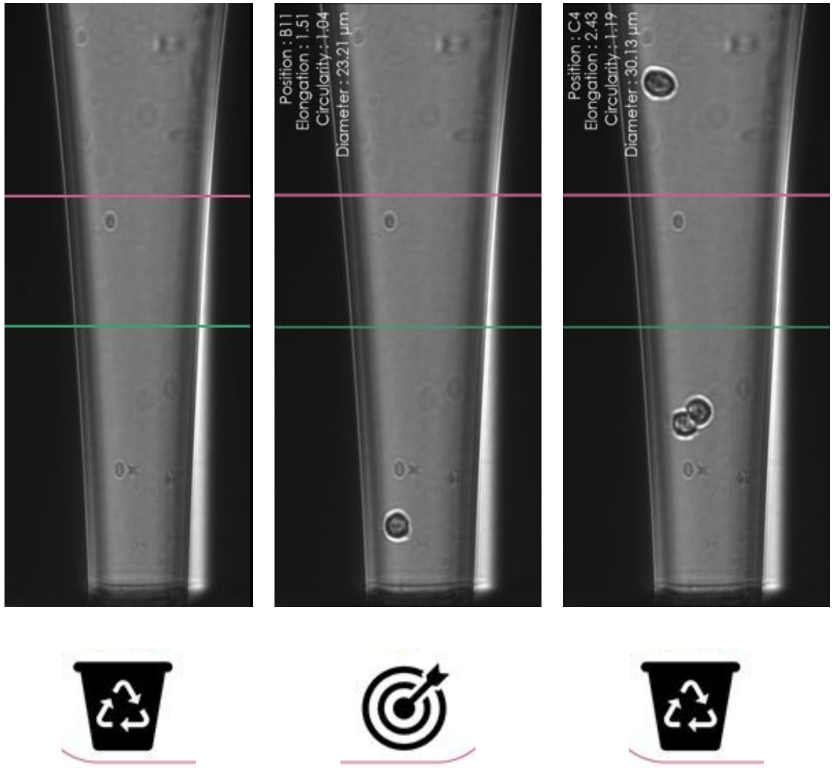
APPENDIX

Some things I muted *(and many, many more)*

- scRNAseq :
  - RNA velocity
  - Protein activity modelization
  - Stemness scoring
  - Variants detection
  - Integration:
    - Multiple samples
  - Multiple omics data
  - All non-droplet methods !
- scEpigenomics : quite everything !
- Other single cell technologies :
  - Genomic :
    - Long reads
  - Non-genomic :
    - Imagery
- Other : ERCCs, PDX, ...

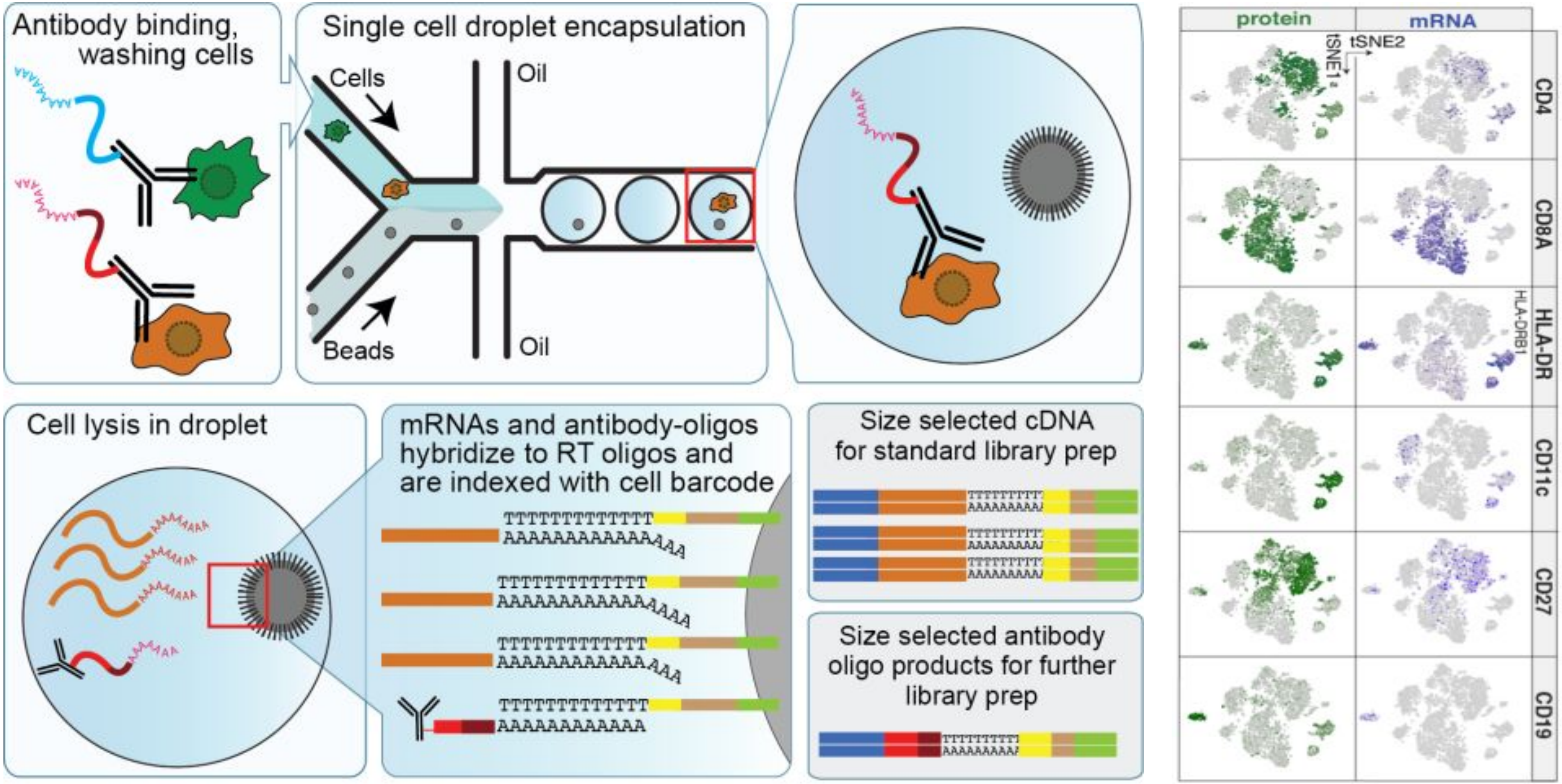
Alternative isolation method : cellenion
IBSCI™ (**I**mage **B**ased **S**ingle **C**ell **I**solation)

- Capillary real-time **video** recording :
  - Cell or no cell ?
  - More than 1 cell ?
  - Cell size ?
- Acoustic dispersion (more gentle)
- Middle scale :
  - Plate-based
  - Up to 1536 cells
- Cell recovery rate over 95%
- Open platform
  - Scalable, compatible
  - Custom reaction kits



CITE-seq
(scRNAseq + proteins)

Cellular Indexing of Transcriptomes and Epitopes by Sequencing

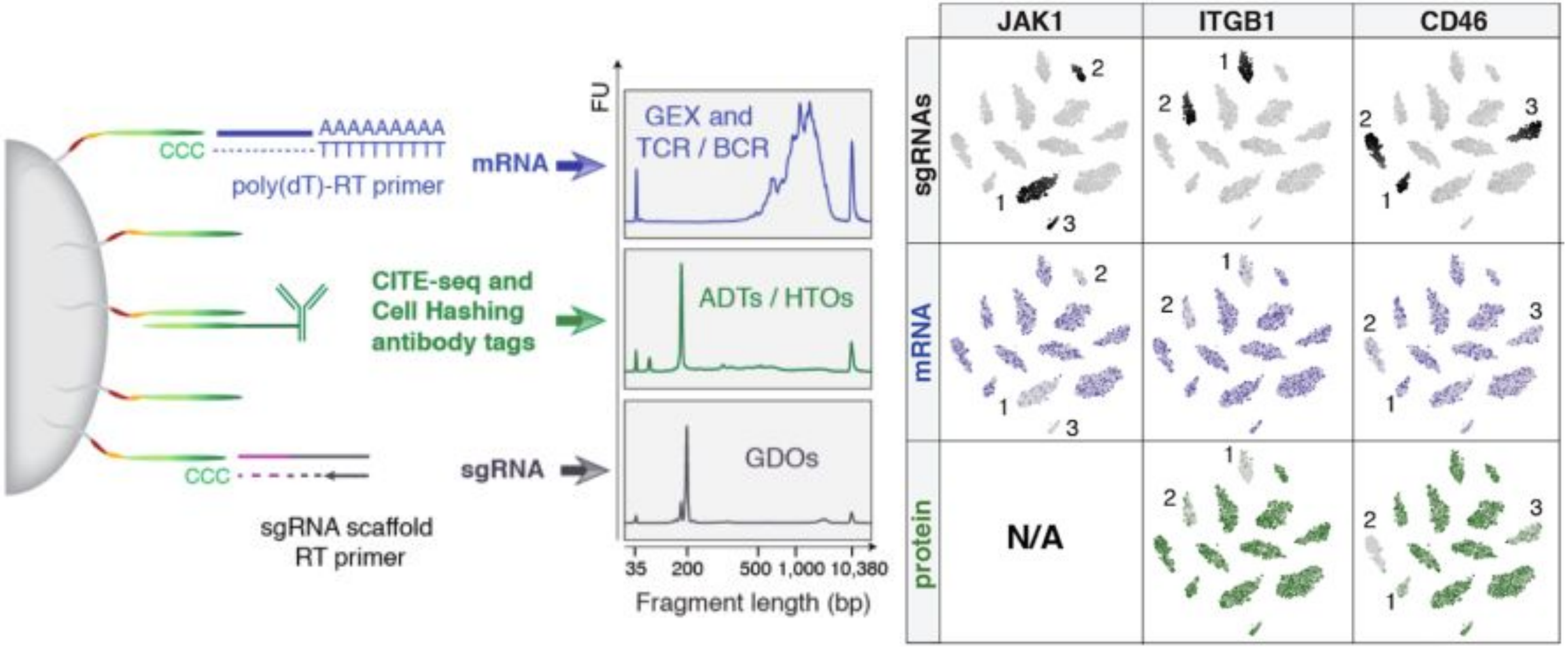


Wikipedia / citeseq.com

ECCITE-seq

(scRNAseq + proteins + CRISPR gRNA)

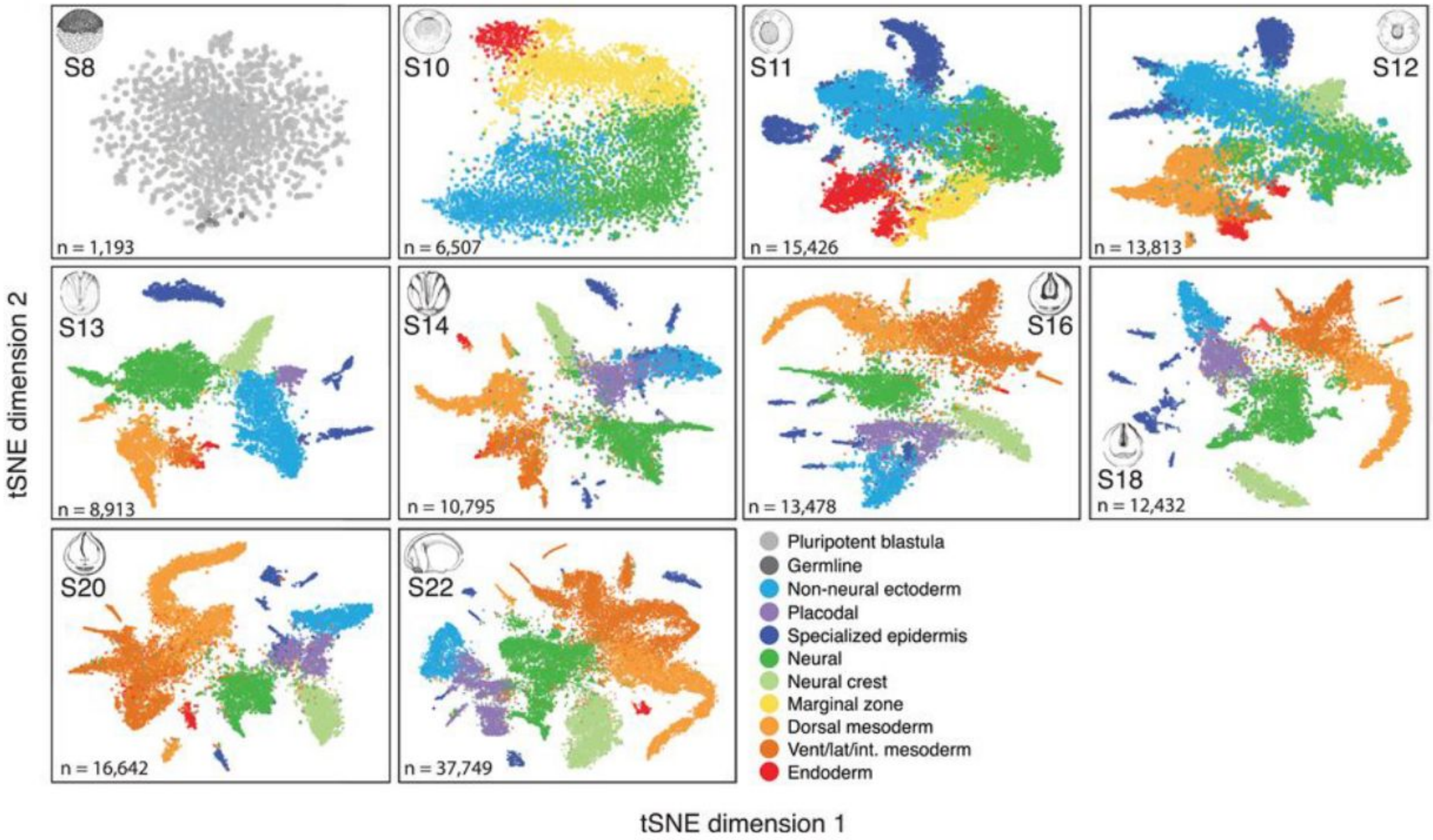
Extended **C**RISPR-compatible **C**ellular **I**ndexing of **T**ranscriptomes and **E**pitopes by **Seq**uencing (5')



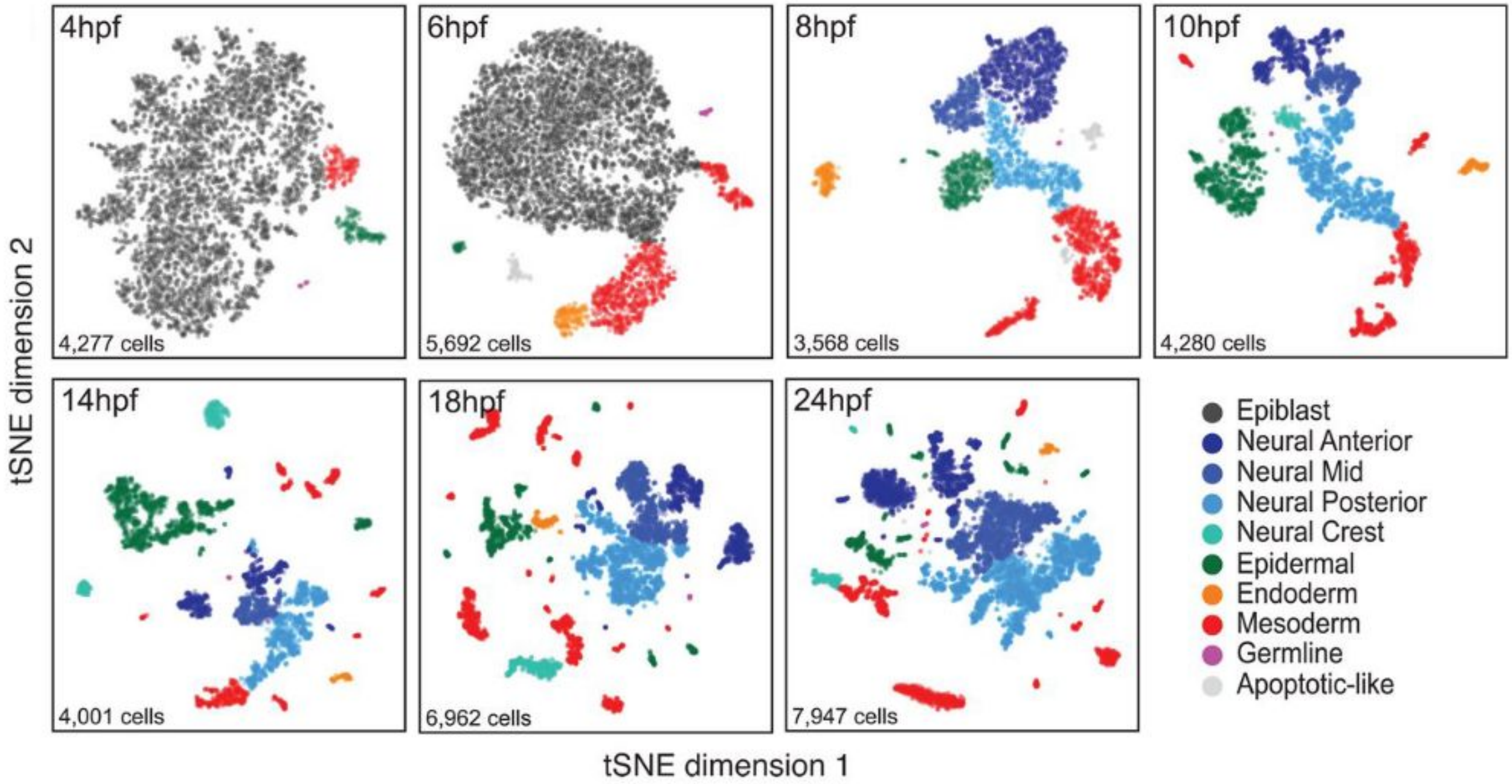
Wikipedia / citeseq.com

Some sweets

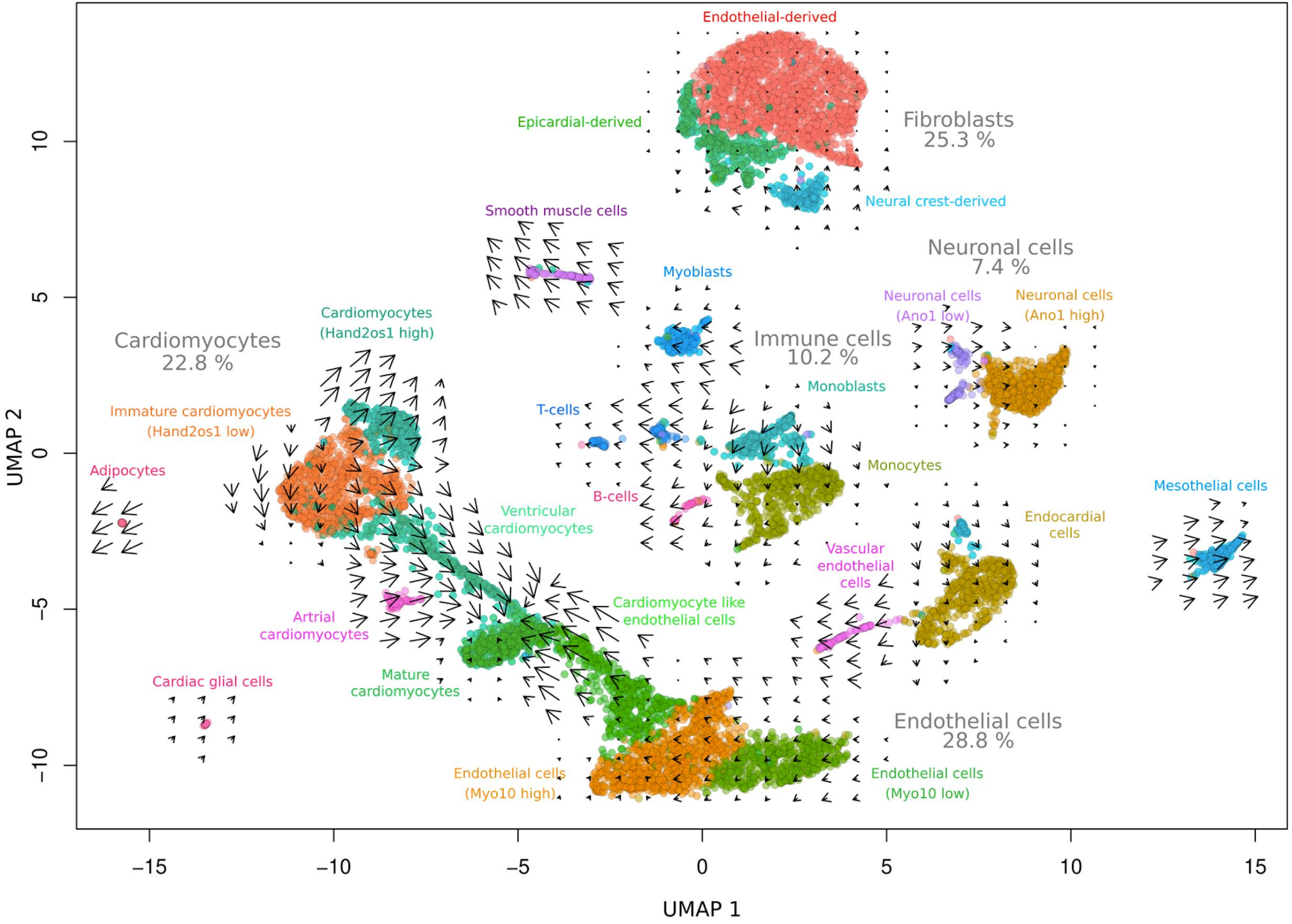
Xenopus embryo development



Zebrafish embryo development

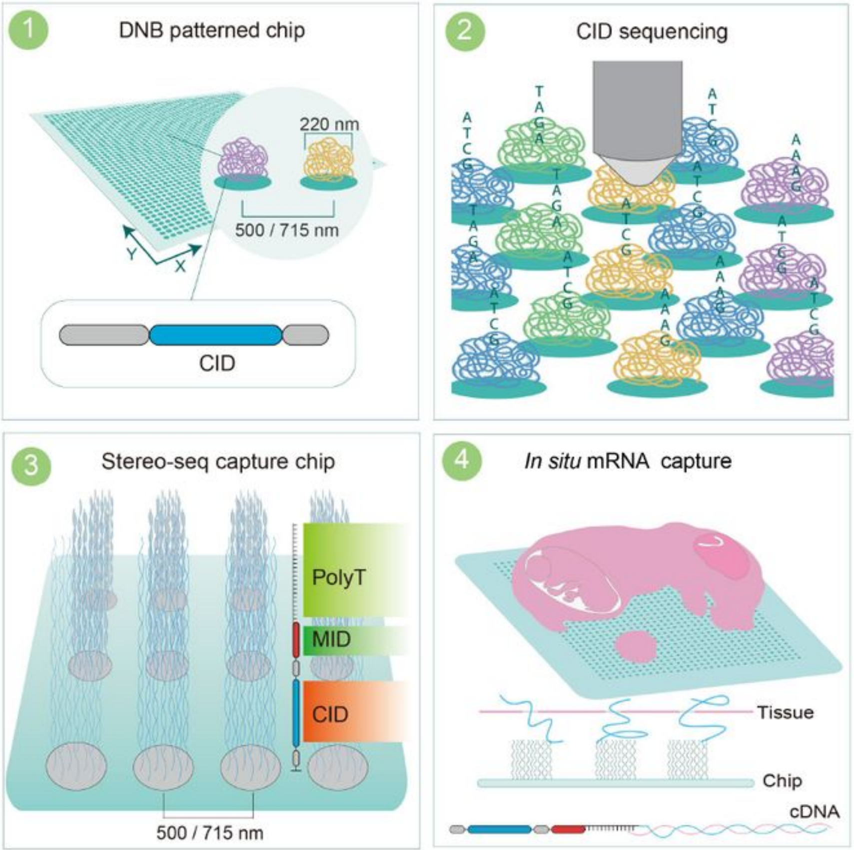


Entire mouse heart : expression & velocity



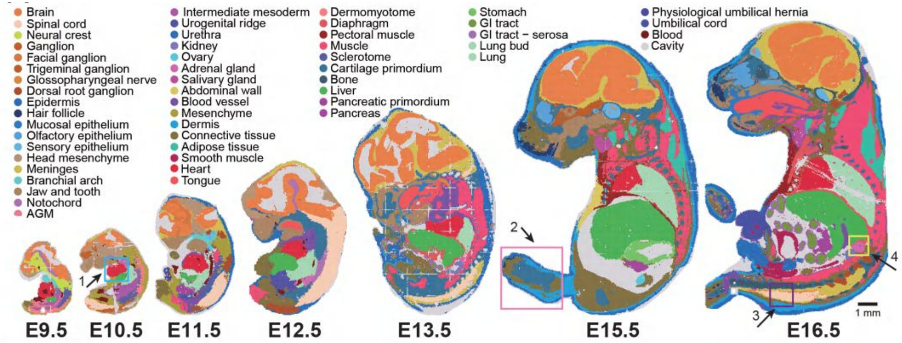
https://doi.org/10.3390/cells9020318

“StereoSeq” : Ultra HD spatial long reads

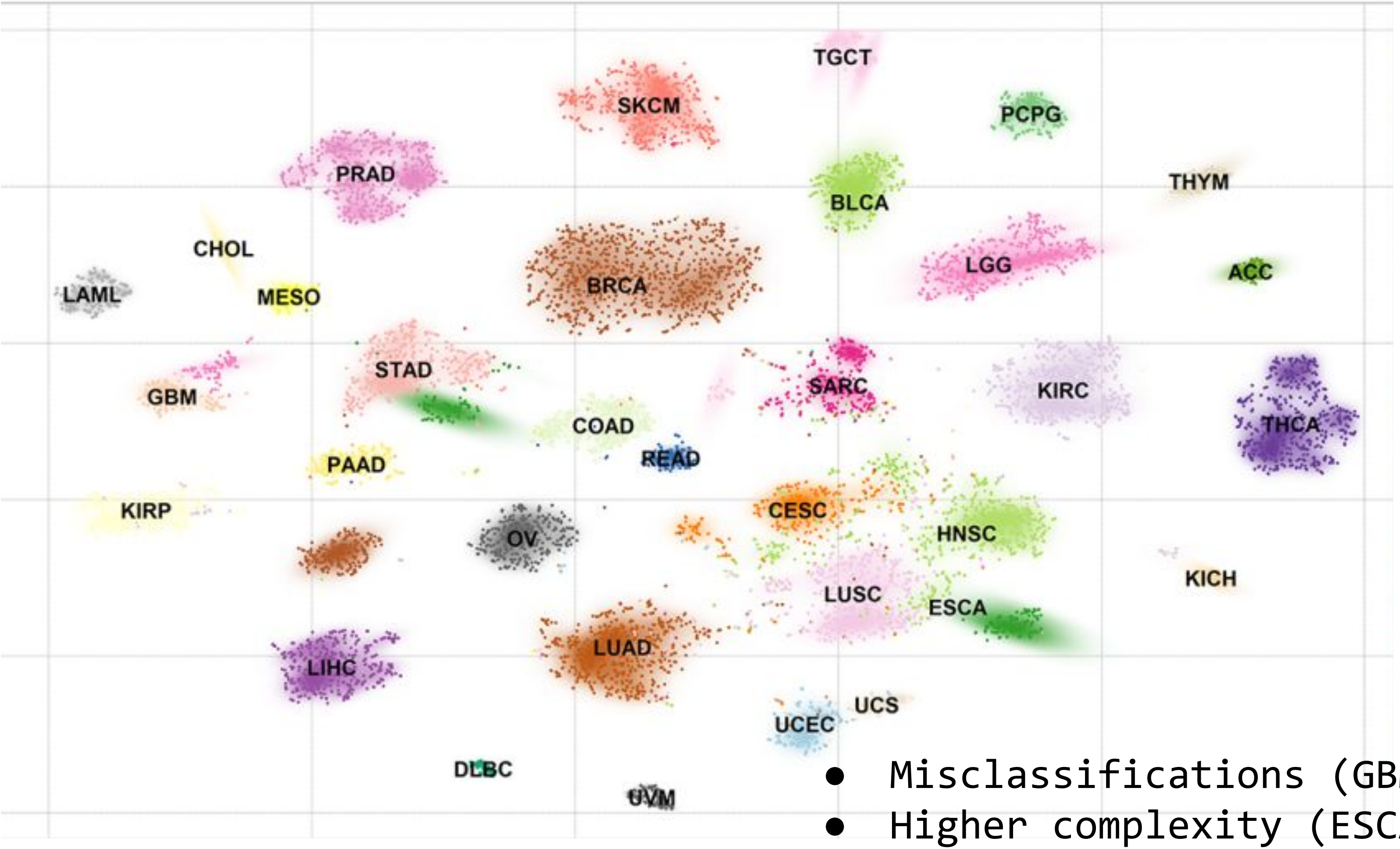


- **Based upon “DNB” (DNA Balls) arrays by MGI/BGI**
- **Claimed resolution of 500-715 nm**
- **Active surface of 200 mm²**
- **Performed a developmental analysis of FULL mice embryos !**

Cheng et al, BiorXiv 2021.01

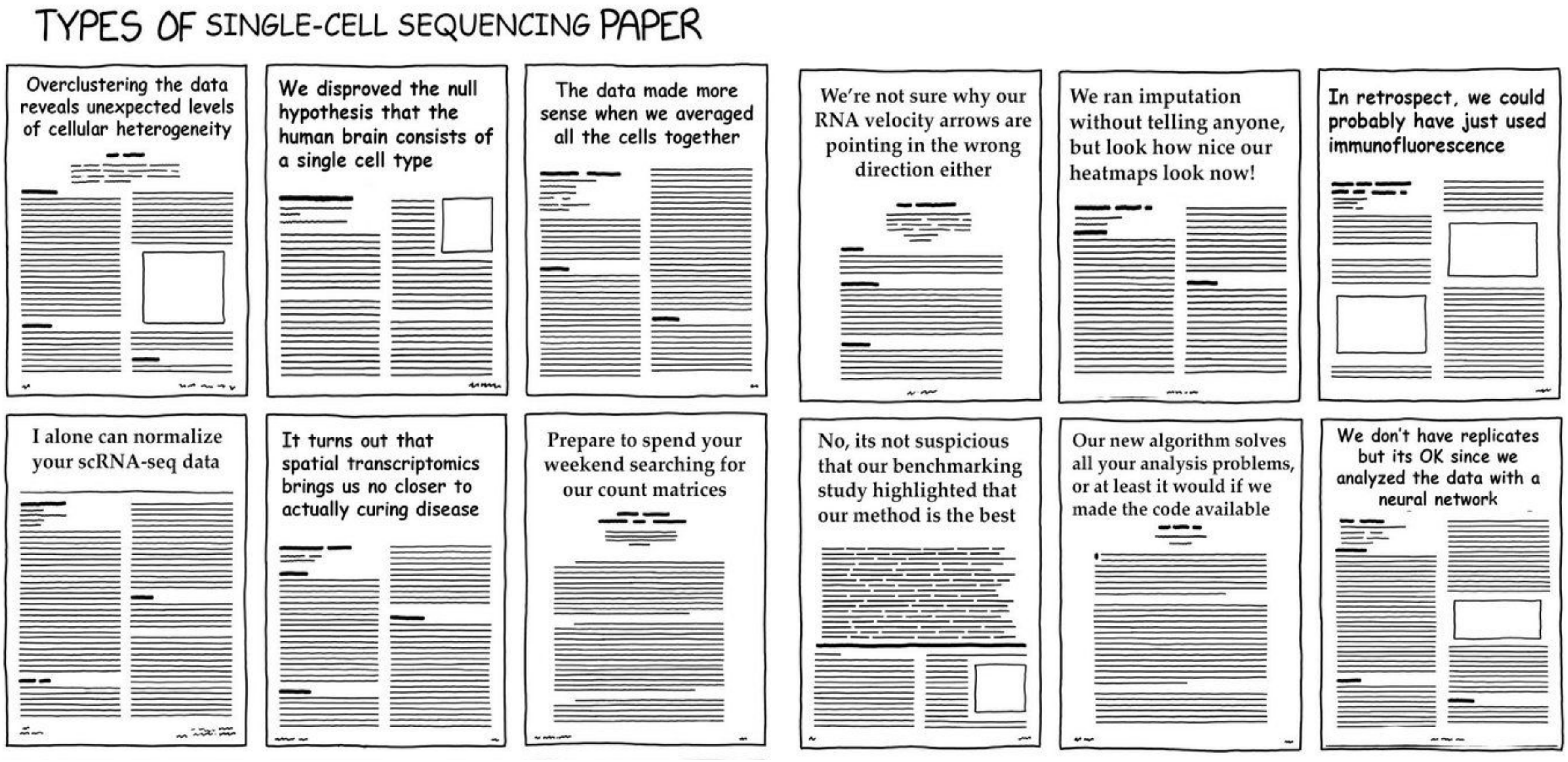


t-SNE of the whole TCGA project (*not SC*)



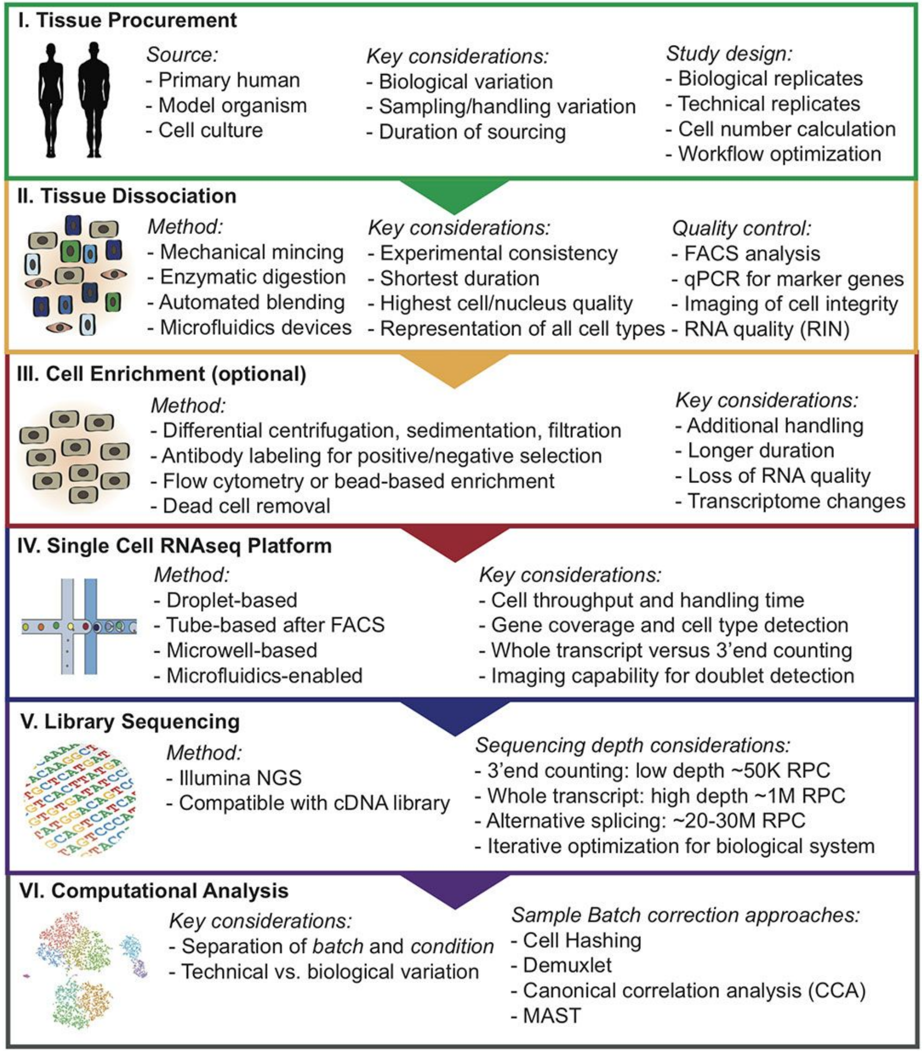
- Misclassifications (GBM-LGG)
- Higher complexity (ESCA-STAD)

Single cell results and the community



tSNE / uMAP plots *are art* !

A little single scRNAseq cheatsheet



Nguyen et al., Front. Cell Dev. Biol. (2019)