

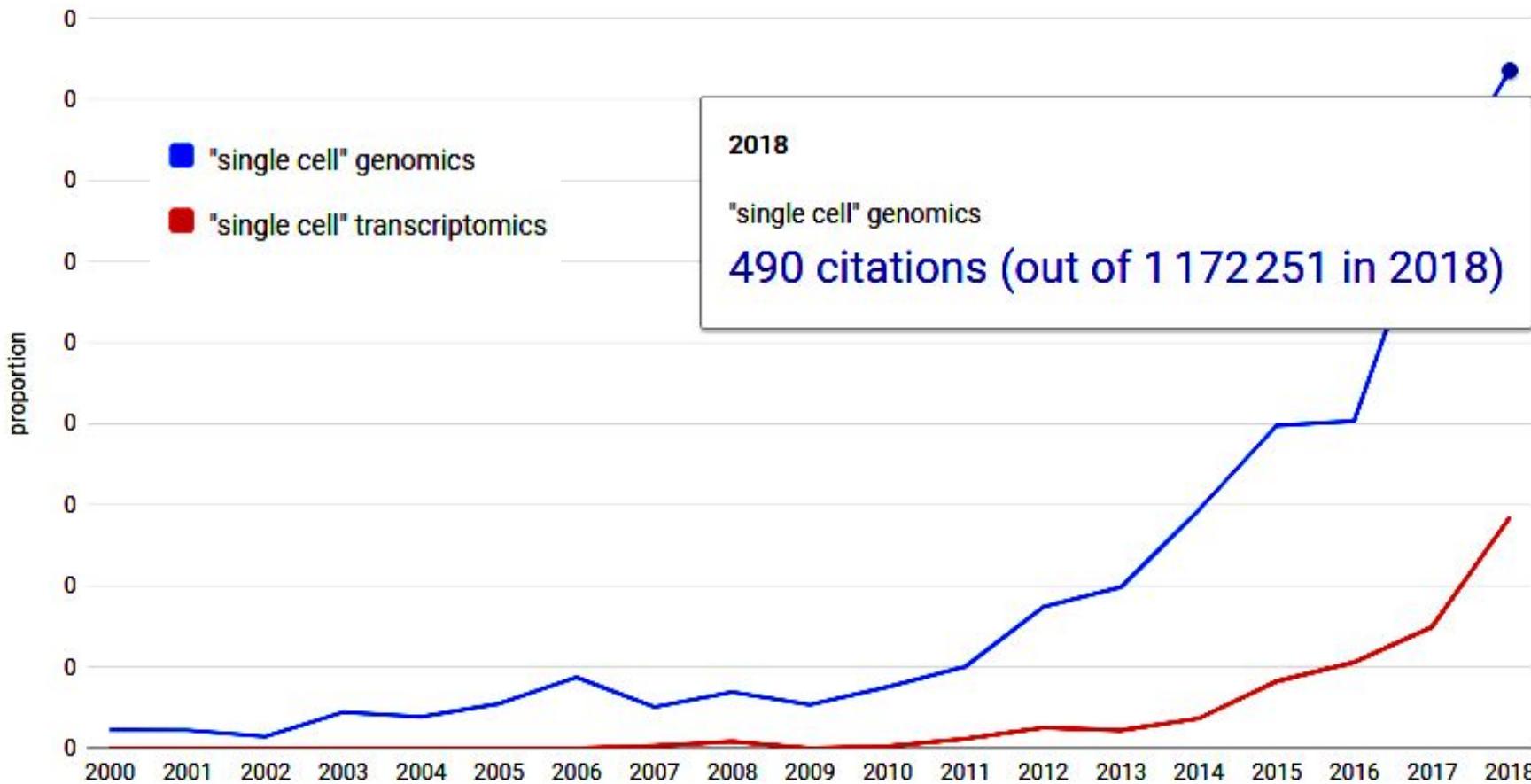
An Introduction to Single-Cell Genomics

Bastien JOB
INSERM / Gustave Roussy

École de bioinformatique AVIESAN-IFB 2019

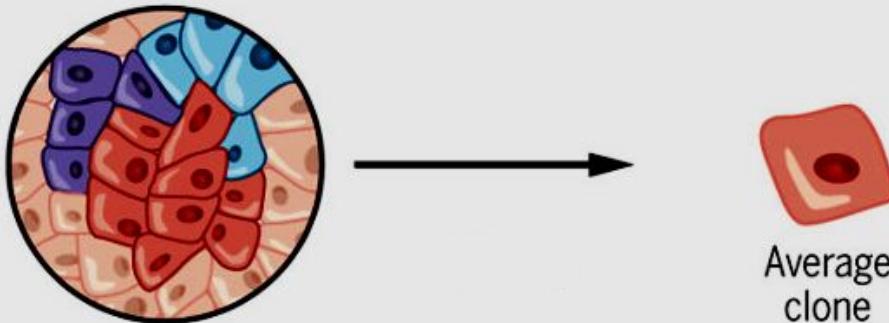
So you say you've heard about single cell ?

Single cell in peer-reviewed publications

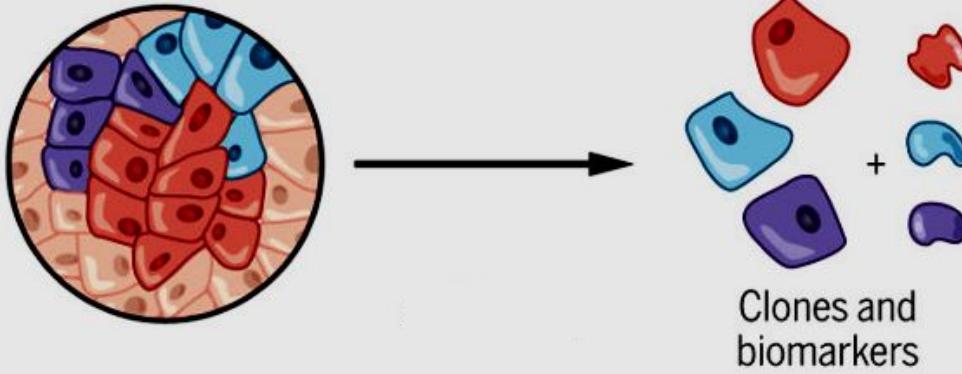


Why so much hype ?

A Bulk analysis



B scRNA analysis

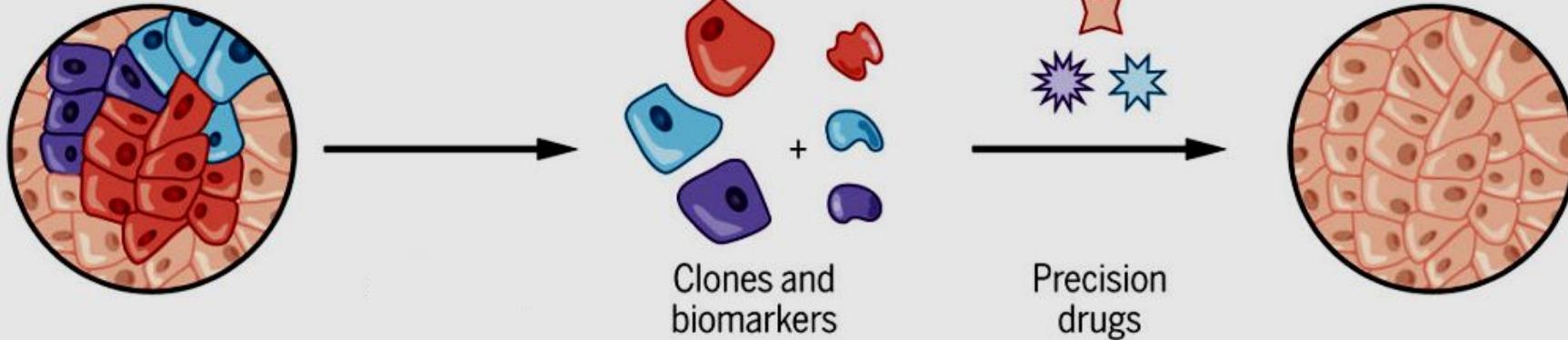


Why so much hype ? (pathology)

A Bulk analysis



B scRNA analysis



Why so much hype ?

Bulk



Single cell

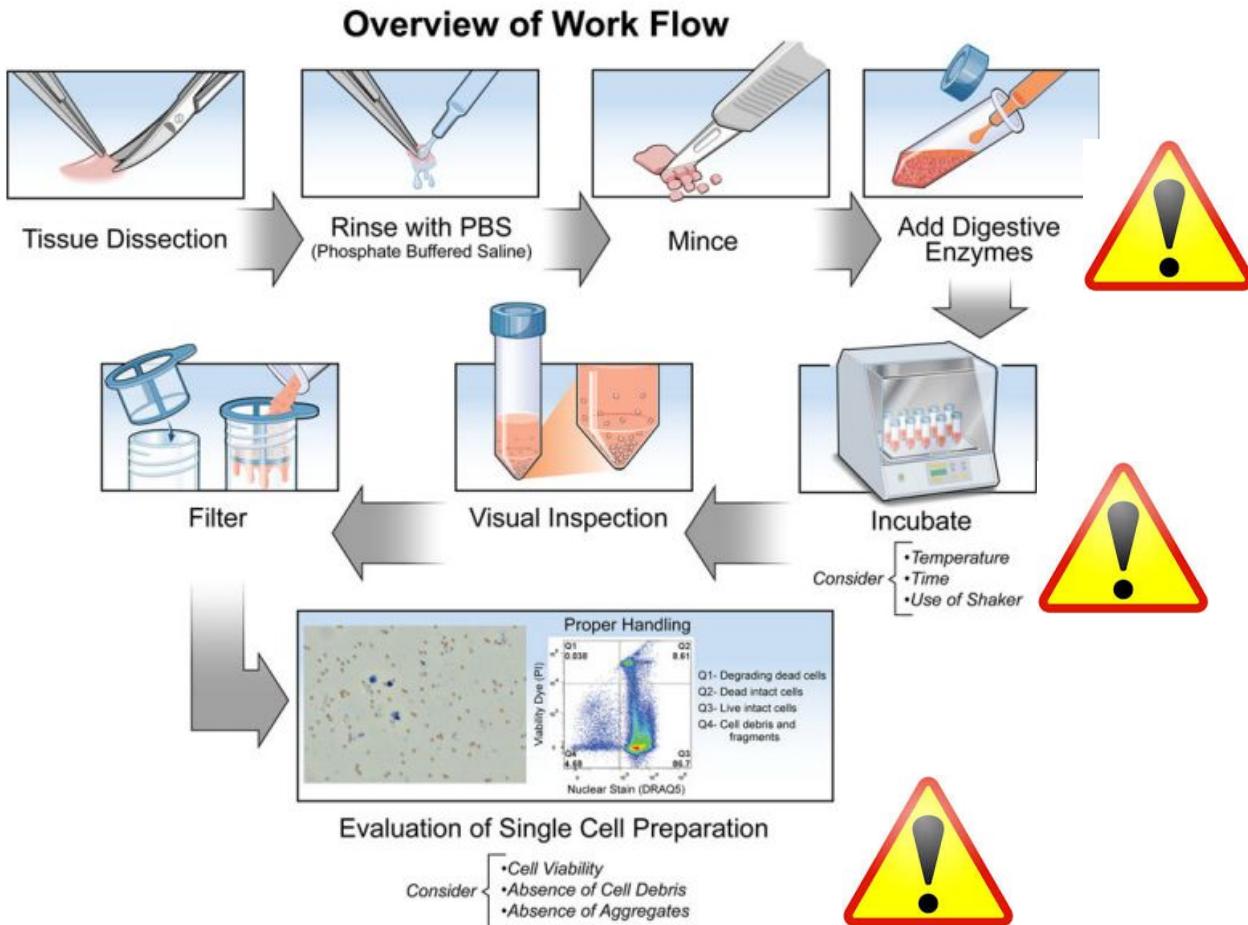


Spatial single cell



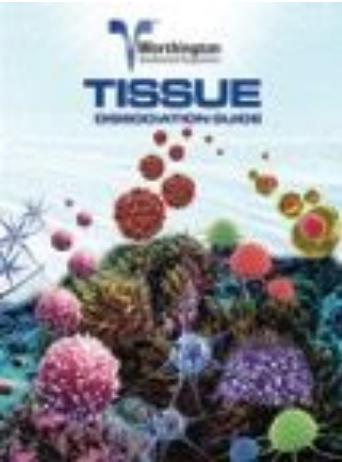
From broad tissue to usable cells

Cells health and dissociation : crucial starting points



Cells health and dissociation : crucial starting points

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

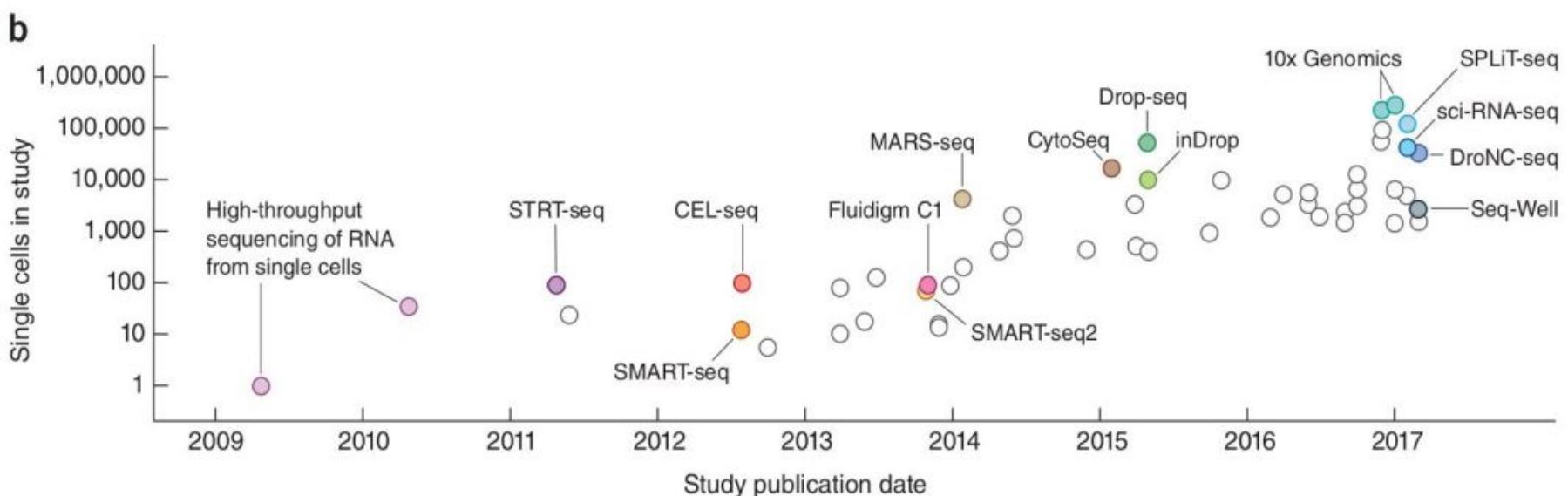
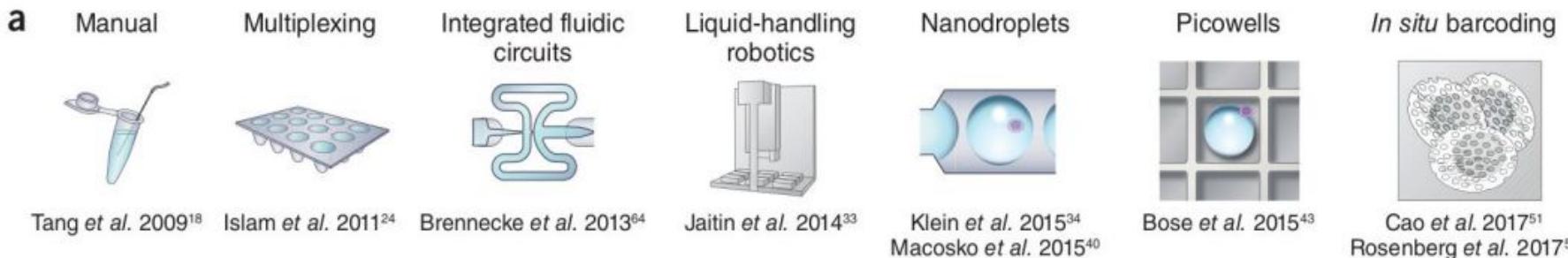
- Tissue Types
 - Epithelial Tissue
 - Connective Tissue
- Dissociating Enzymes
 - Collagenase
 - Trypsin
 - Elastase
 - Hyaluronidase
 - Papain
 - Chymotrypsin
 - Deoxyribonuclease I
 - Neutral Protease (Dispase)
 - Trypsin Inhibitor
 - Animal Origin Free (AOF) Enzymes
 - Celase® GMP

III. Cell Isolation Techniques

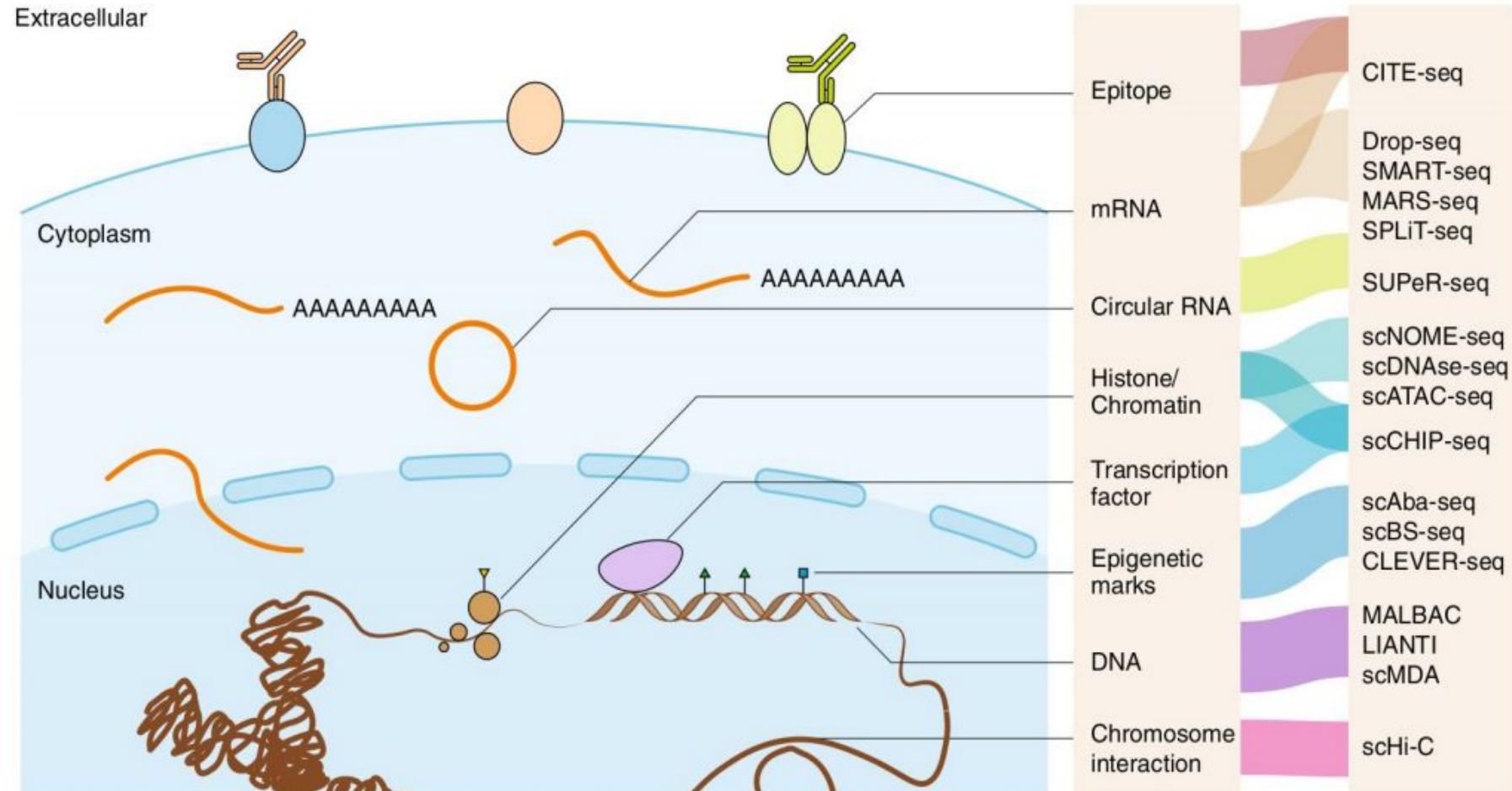
- Methods & Materials
 - Working With Enzymes
 - Basic Primary Cell Isolation
 - Equilibration with 95%O₂:5%CO₂
 - Trituration
 - Enzymatic Cell Harvesting
 - Cell Adhesion and Harvesting
 - Trypsin for Cell Harvesting
 - Cell Release Procedure
- Optimization Techniques
 - General Guidelines
 - Optimization Strategy
 - Cell Quantitation
 - Measure of Viability

IV. Use-Tested Cell Isolation Systems

Cells separation : several technologies over the last decade

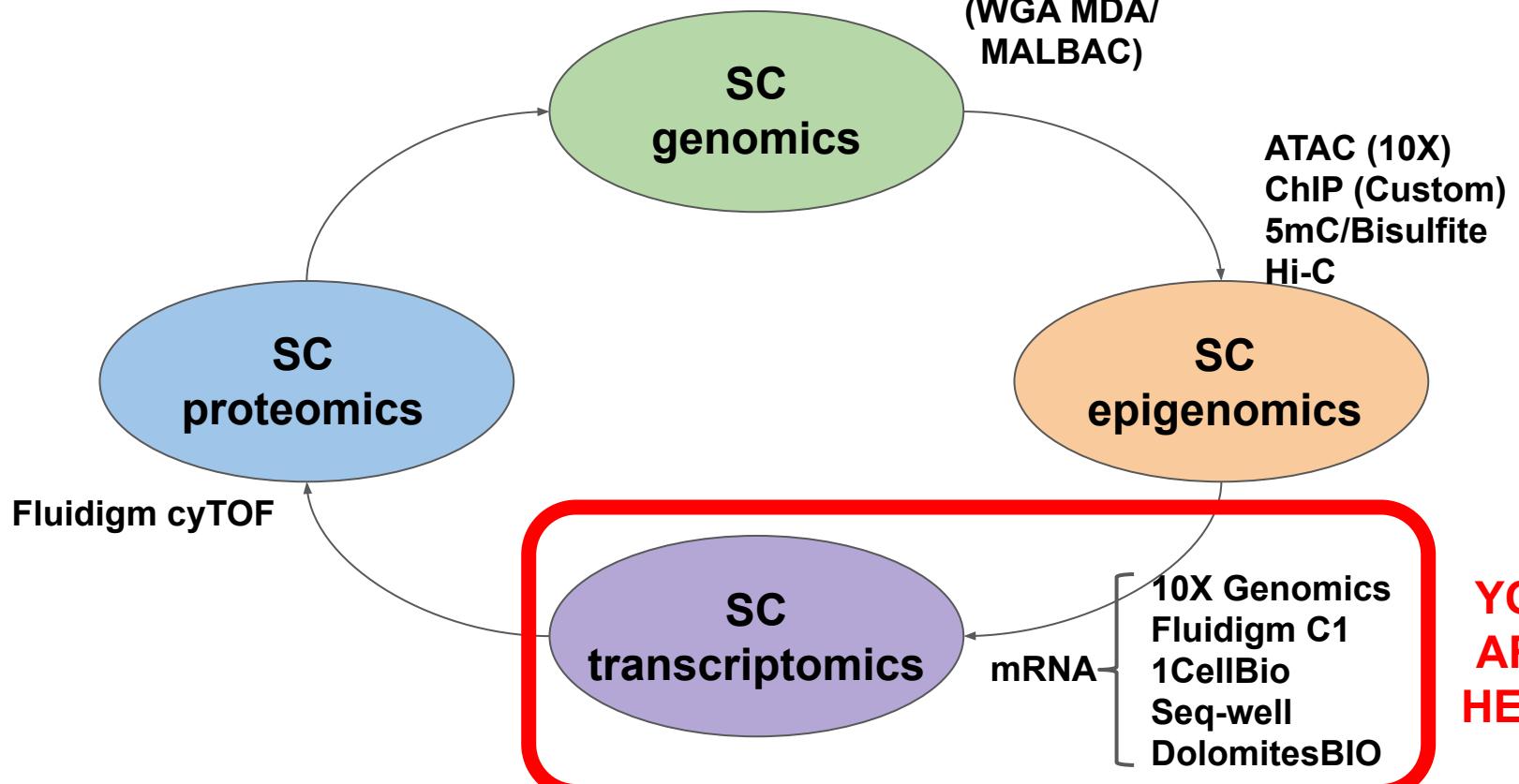


Several protocols for several purposes

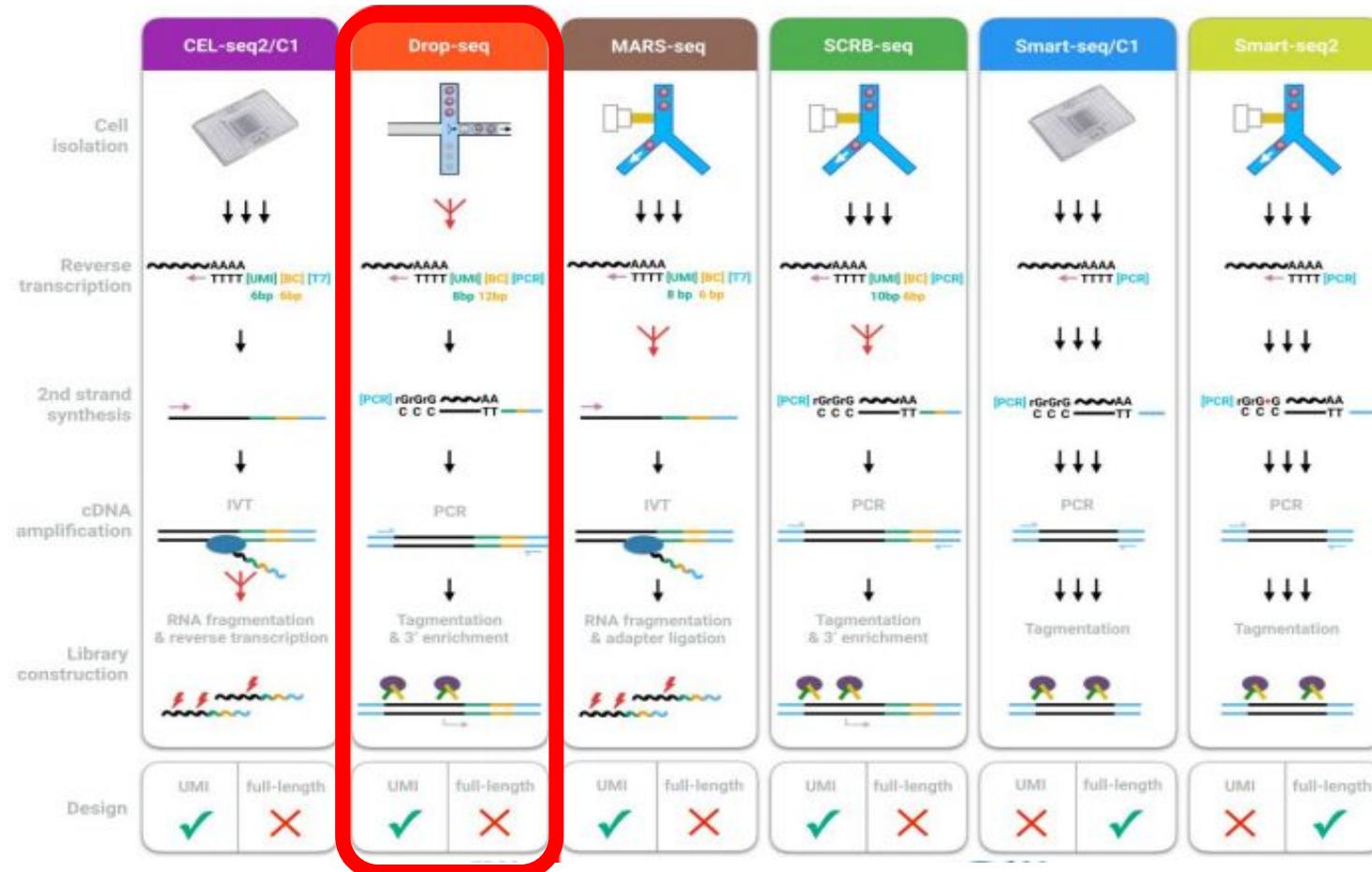


From isolated cells to nucleotide sequences

Single Cell RNAseq

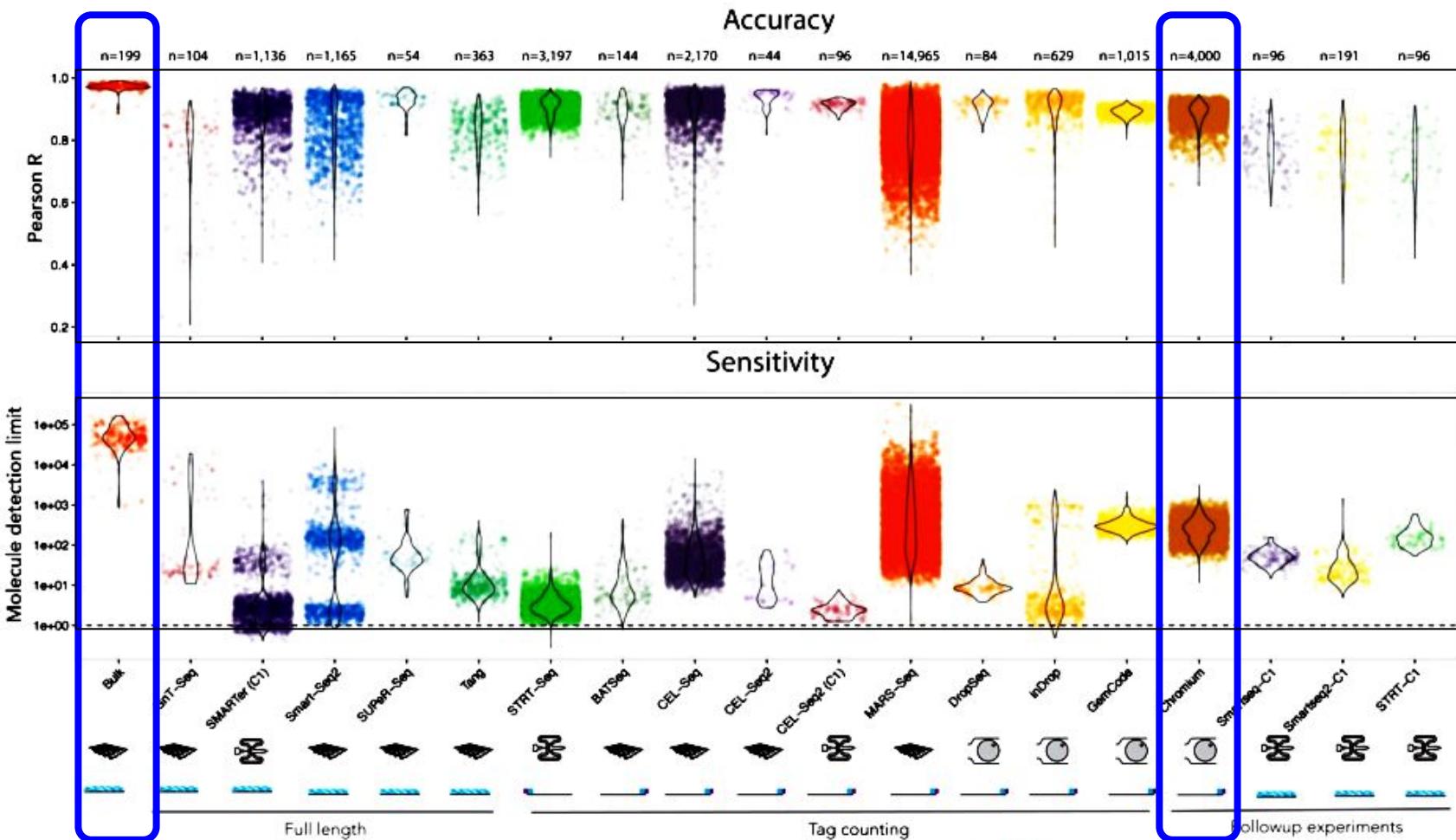


From usable cells to sequences (Drop-seq / 10X)



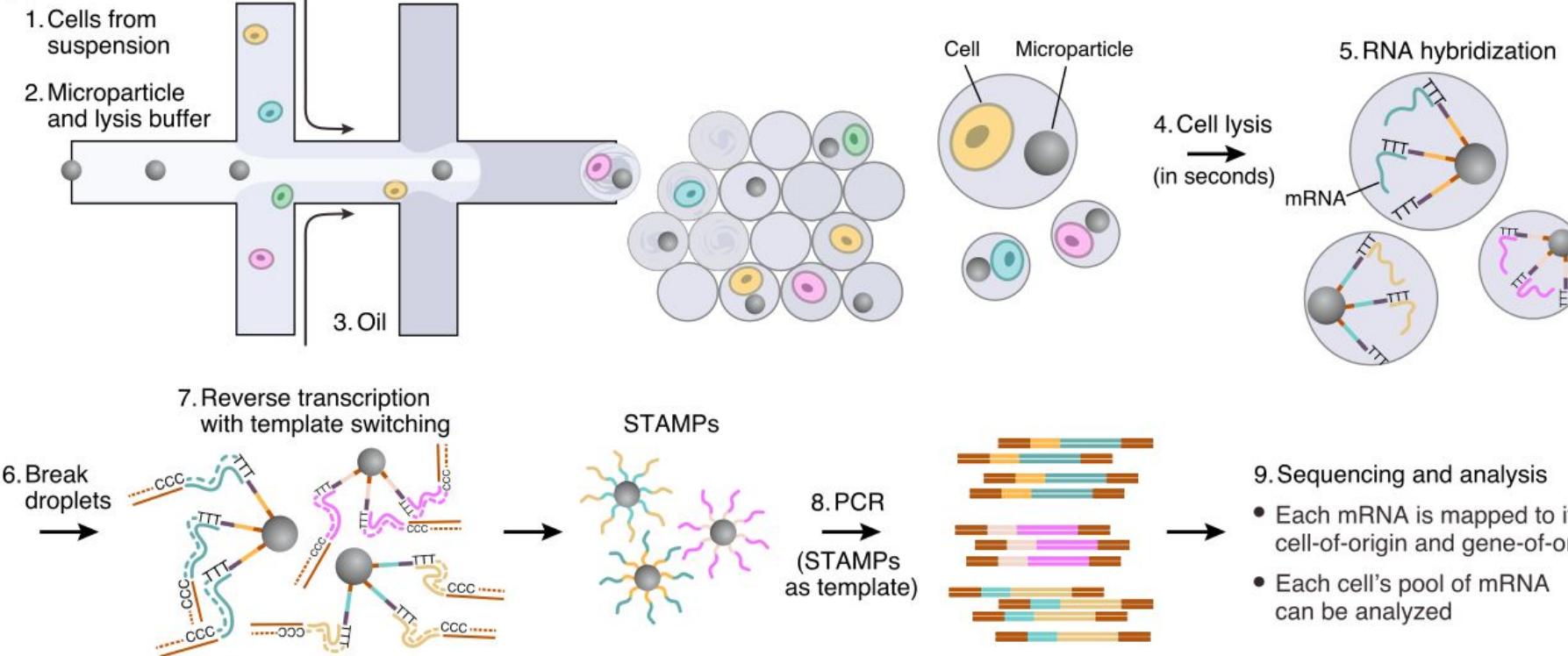
Bulk

10X



Drop-seq

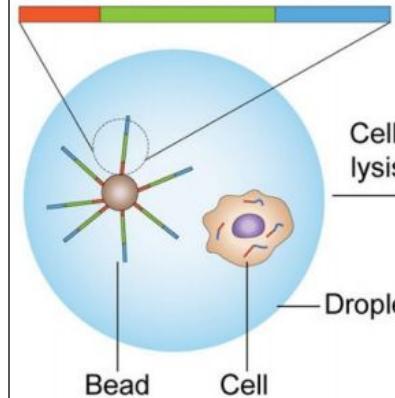
A



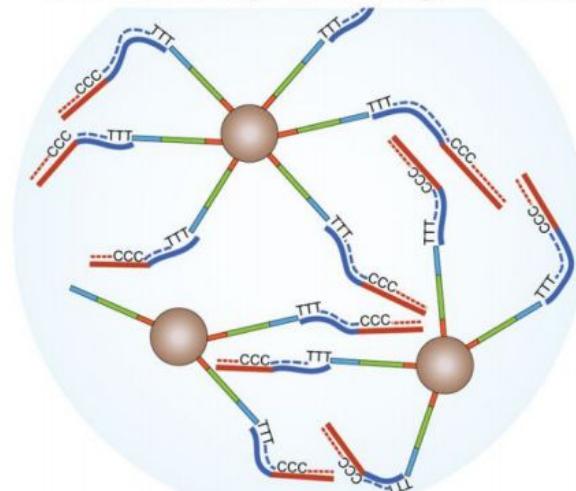
10X Chromium (3')

Structure of the barcode primer bead

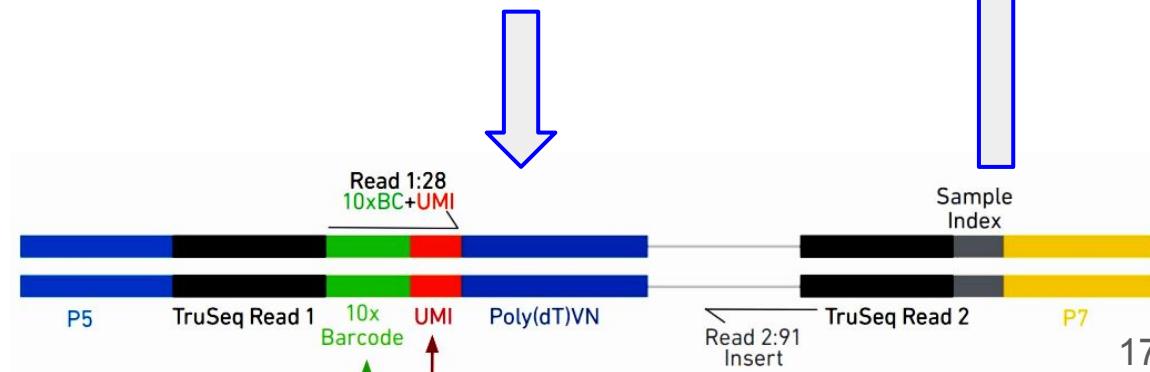
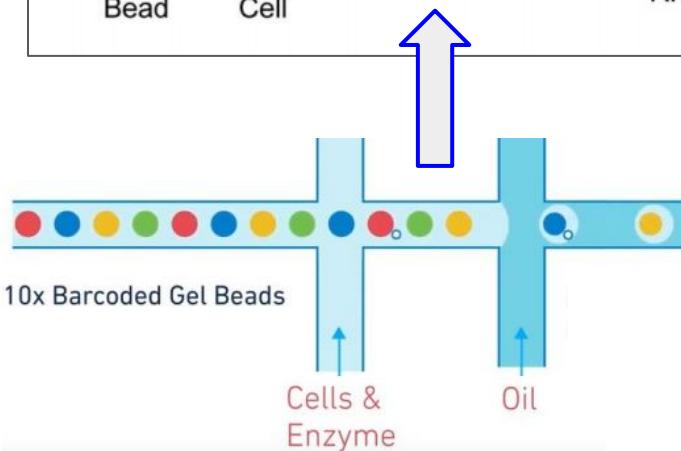
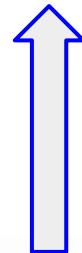
PCR
handle Cell barcode UMI



Reverse transcription with template switching



Sequencing



From nucleotide sequences (reads) to count matrices



Reads QC

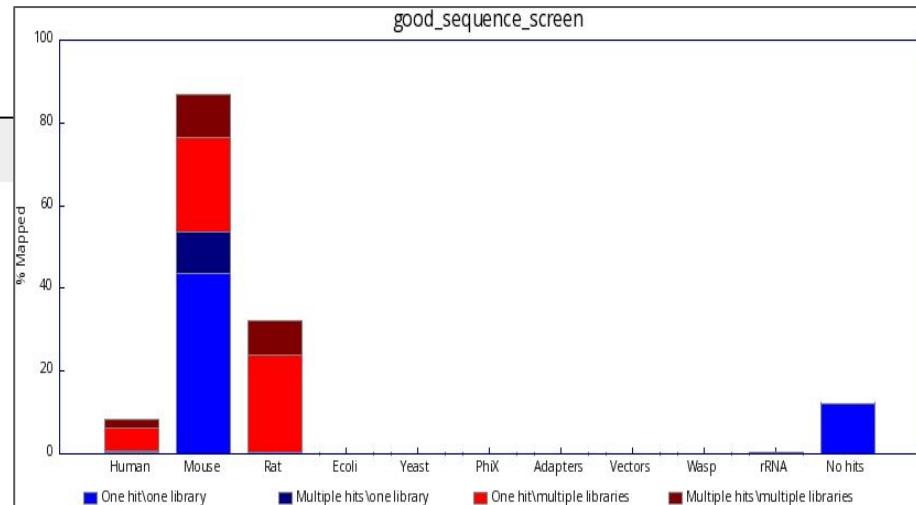
FastQC Report

Summary

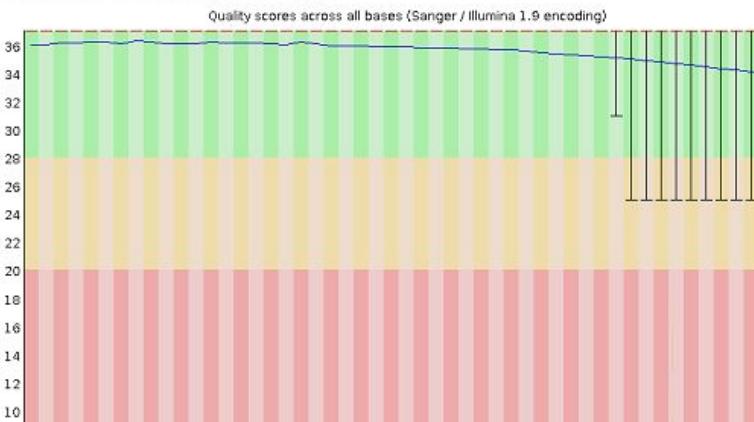
- ✓ Basic Statistics
- ✓ Per base sequence quality
- ✗ Per tile sequence quality
- ✓ Per sequence quality scores
- ✗ Per base sequence content
- ✓ Per sequence GC content
- ✓ Per base N content
- ✓ Sequence Length Distribution
- ✗ Sequence Duplication Levels
- ✗ Overrepresented sequences
- ✓ Adapter Content

Basic Statistics

Measure	Value
Filename	BC_392_1_529_R2_001.fastq.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	109443265
Sequences flagged as poor quality	0
Sequence length	91
%GC	43



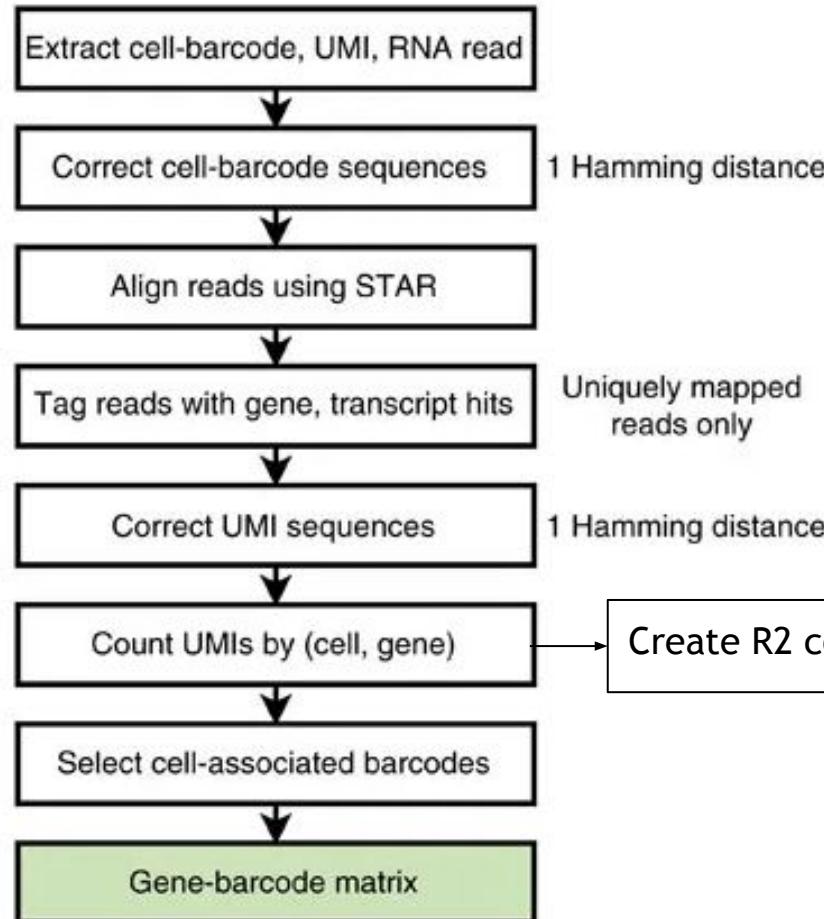
Per base sequence quality



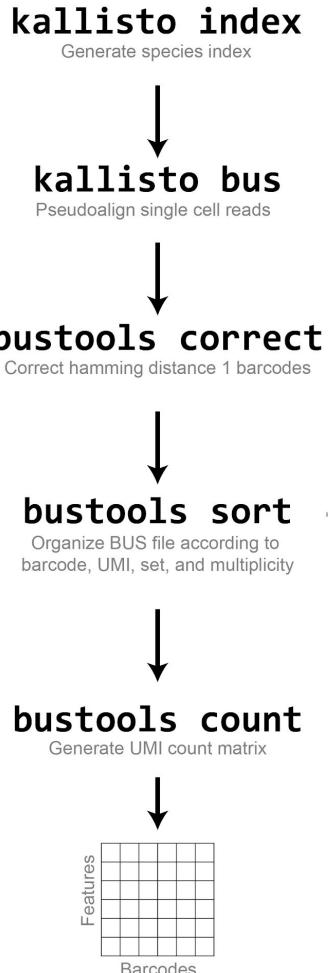
- As usual : FASTQC, FastqScreen, ...
- Special attention to R1 : Cell barcode & UMI !
- Control of the 4 sample libs

Reads processing workflows

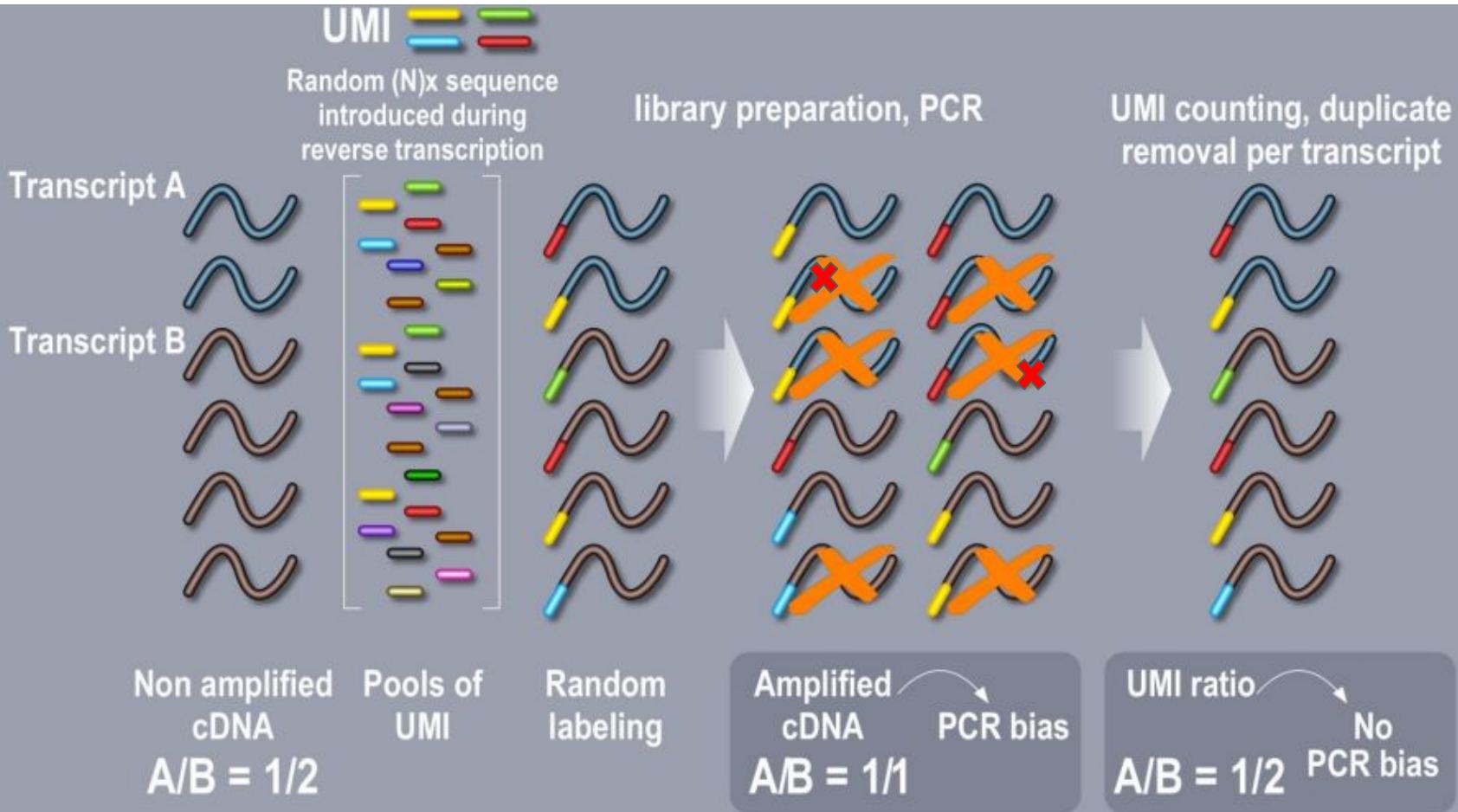
Mapping-based (STAR)



Pseudomapping-based (kallisto bustools)

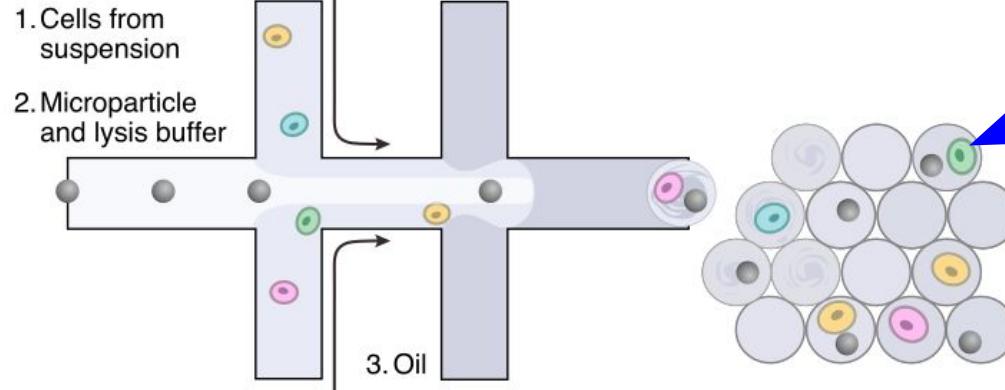


Focus on : Unique Molecule Identifiers



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

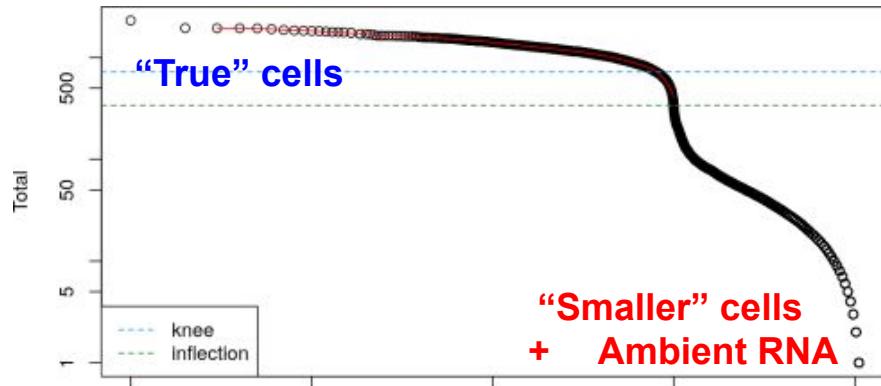
Focus on : Empty droplets filtering



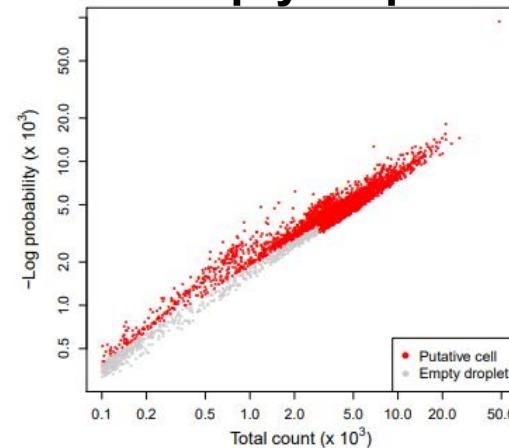
THERE IS RNA HERE
(CELL IN GEM)

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

Kneeplot

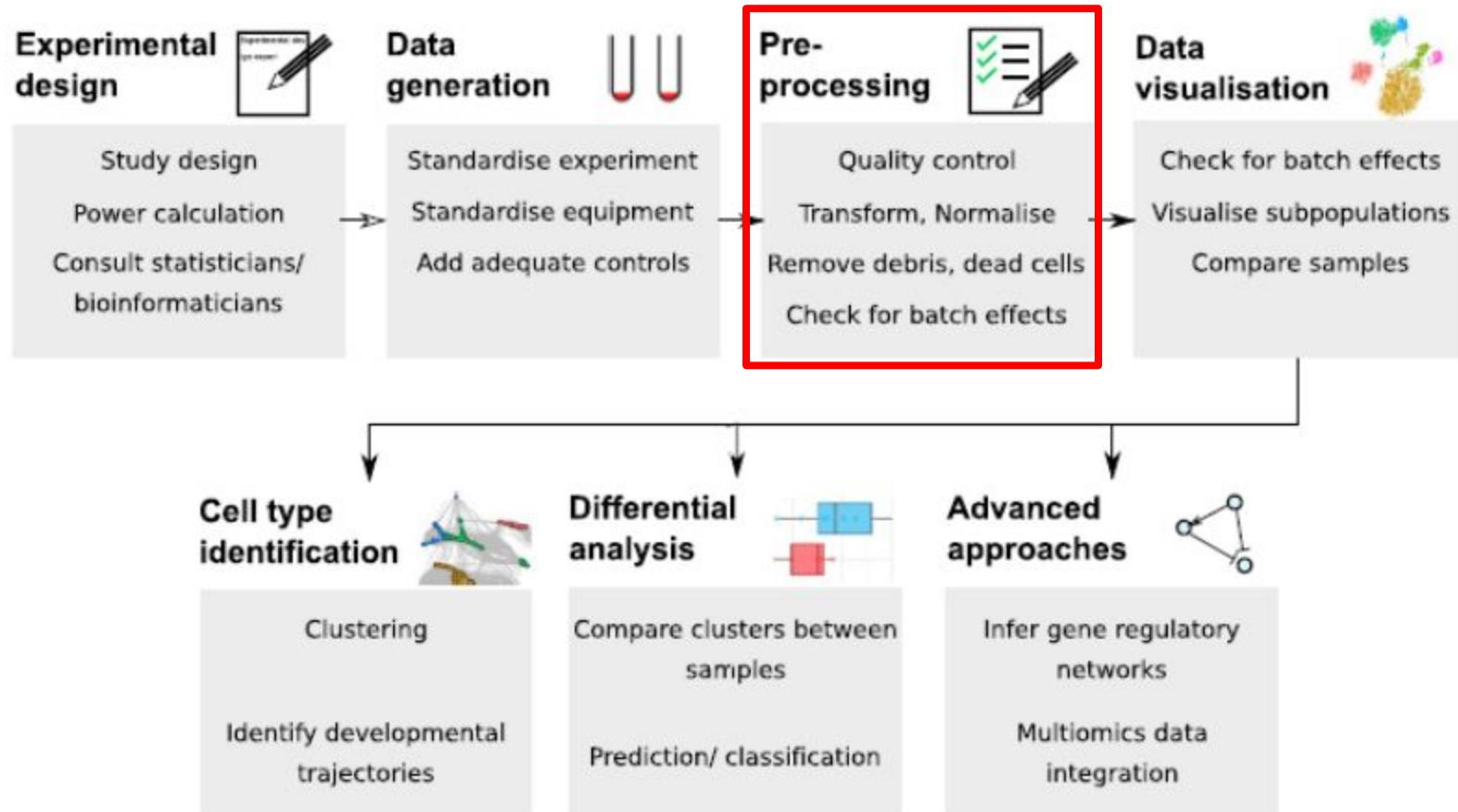


EmptyDrops



From raw counts matrix to normalized data

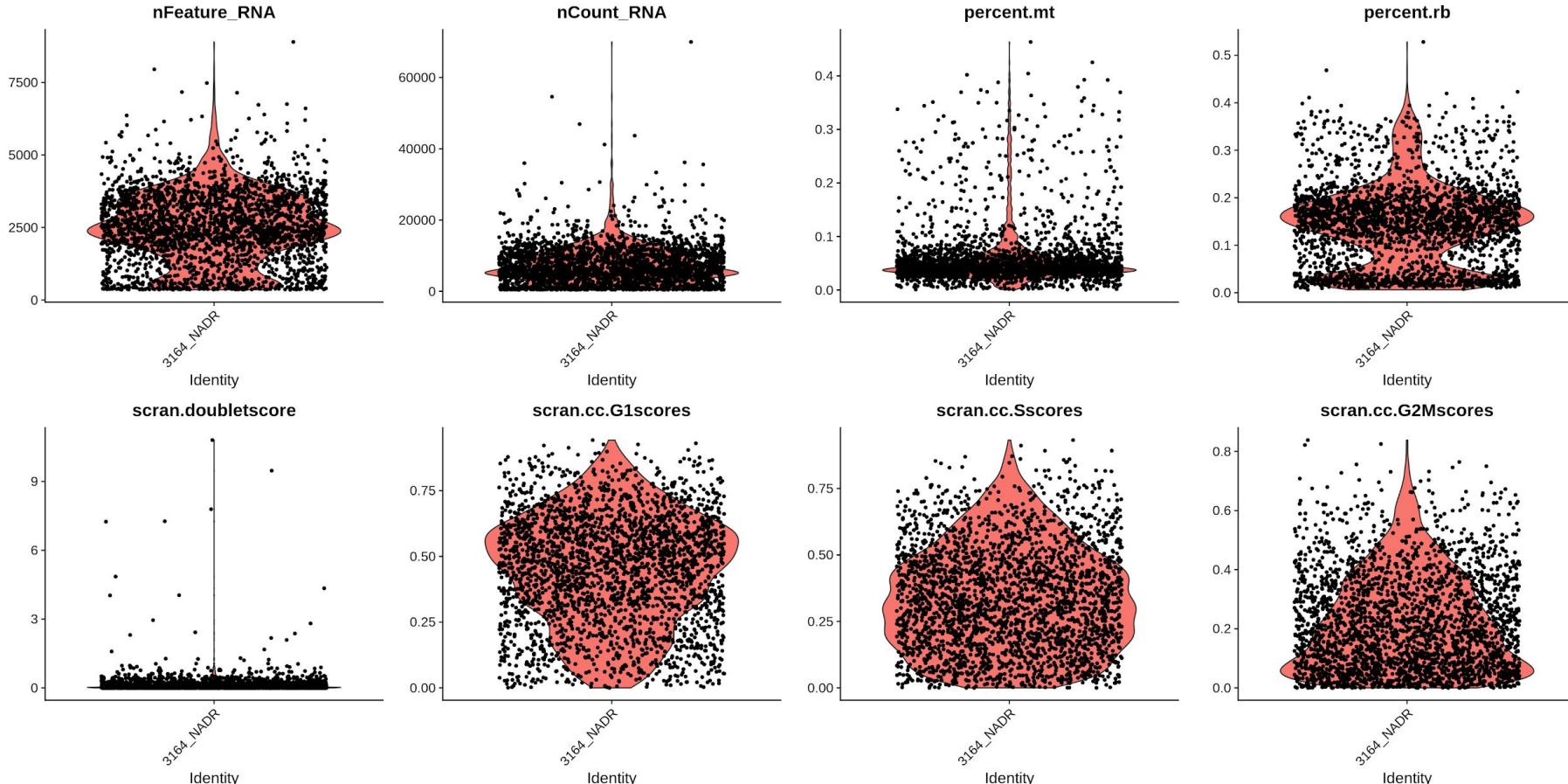
Standard analysis pipeline



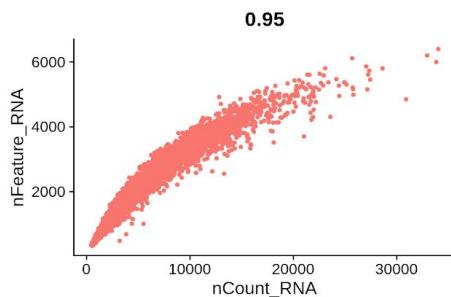
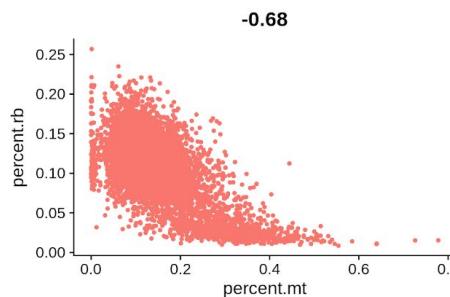
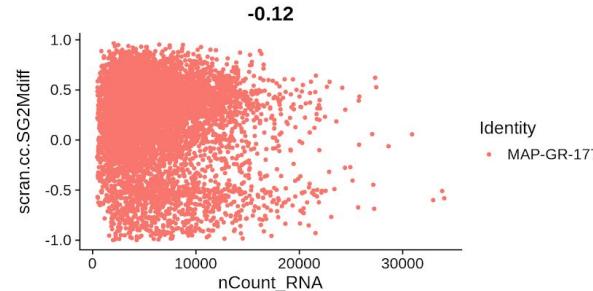
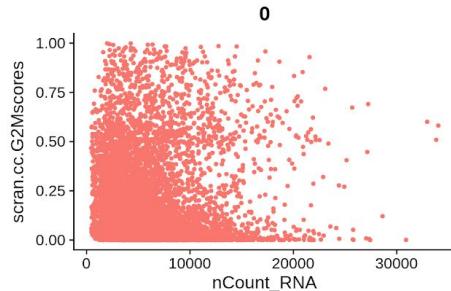
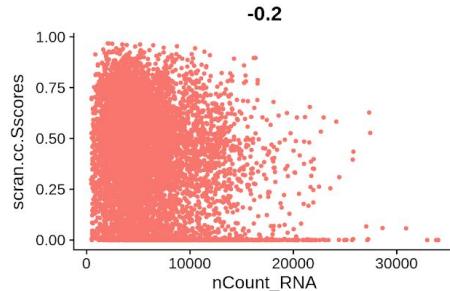
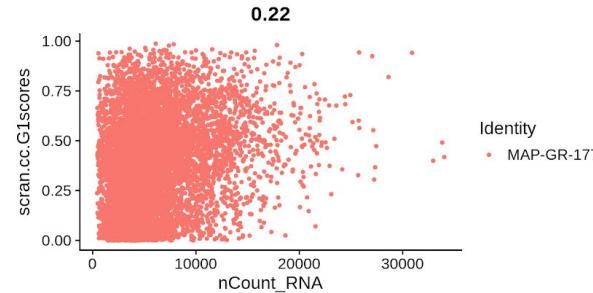
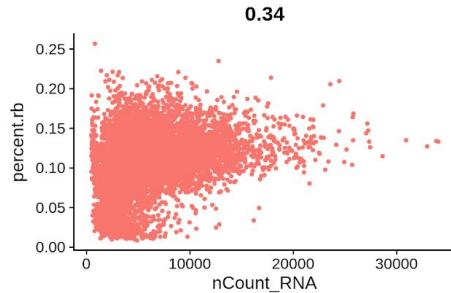
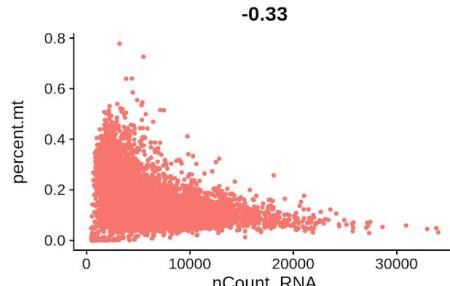
Cell QC considerations

- 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

Cell QC : metrics

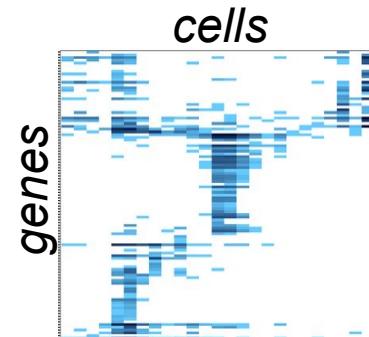
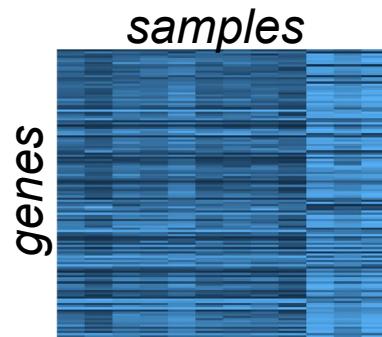


Cell QC : metrics



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)

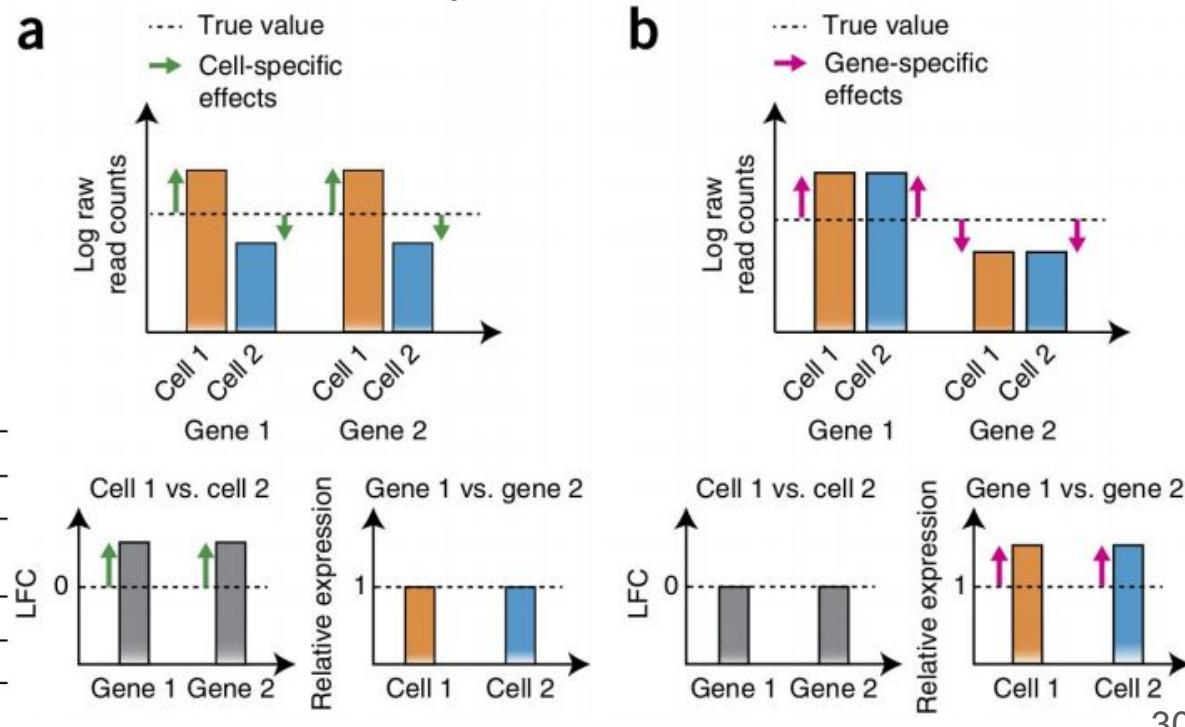


SC MATRIX IS SPARSE ! (ie, mostly filled with zeros)

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



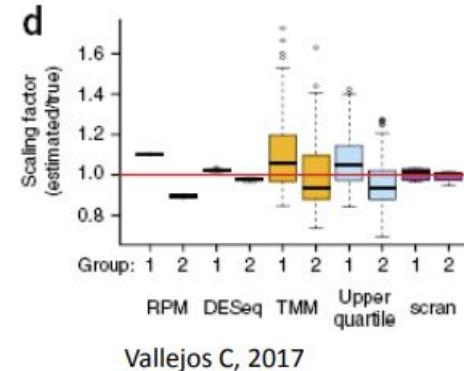
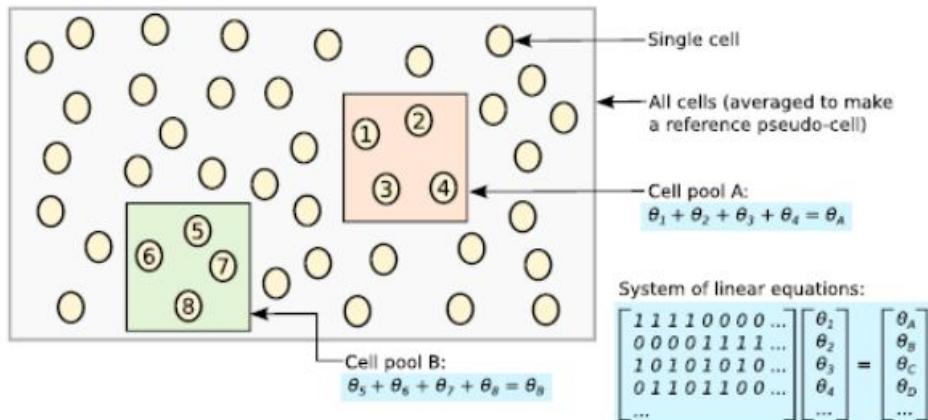
Global scaling : Bulk normalization methods are KO

- 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 $\geq 50\%$ 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

Matrix normalization : scaling by factors

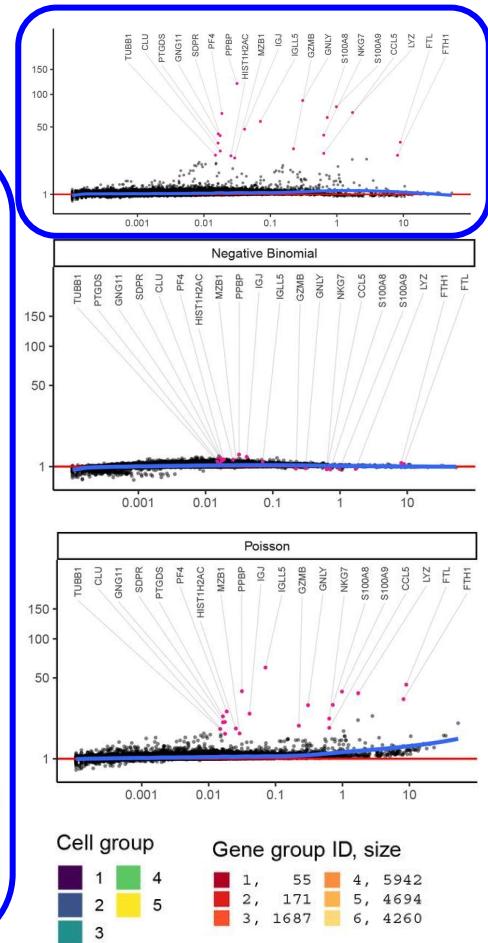
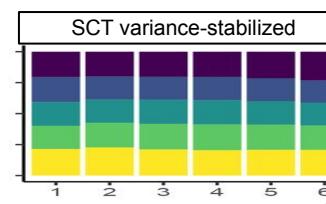
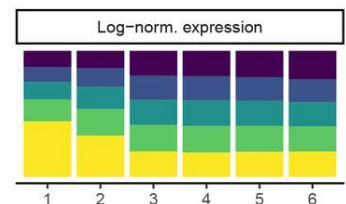
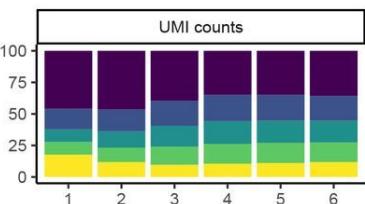
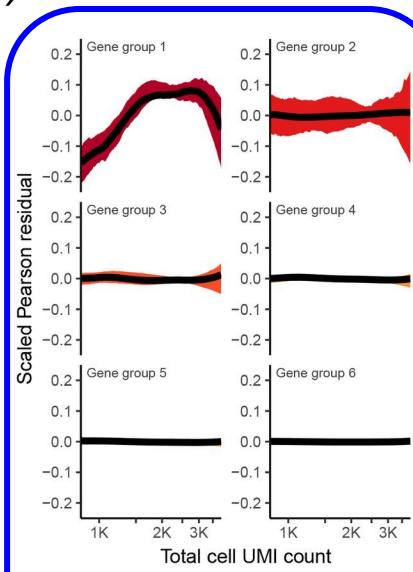
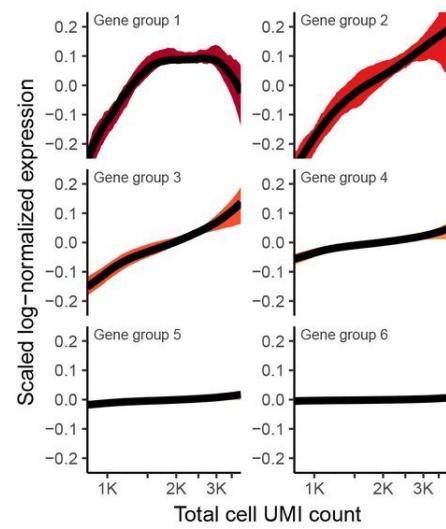
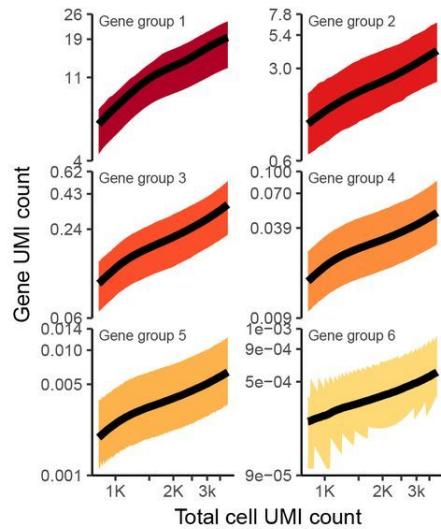
- Alternative method to compute the size factors
- Pool cells to reduce the number of zeros
- Estimate the size factors for the pool
- Repeat many time and use deconvolution to estimate each cell size factor
- Implemented in **scater/scran** packages



Vallejos C, 2017

Matrix normalization : variance stabilization

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

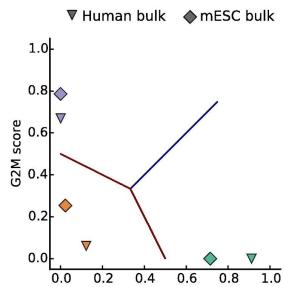


Hafemeister et al, BioRxiv (2019)

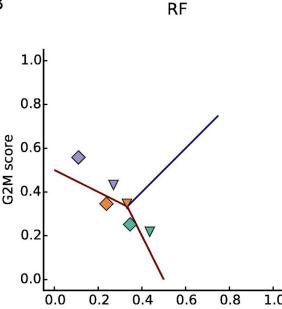
Other factors : Cell cycle phase

- 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*)

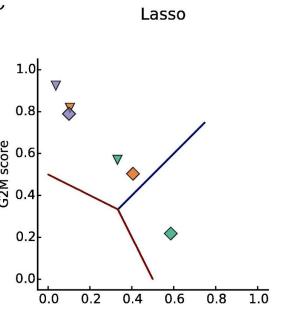
A PCA



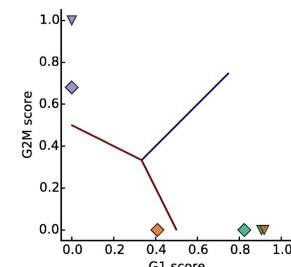
B RF



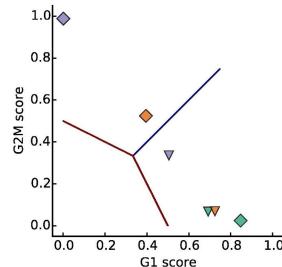
C Lasso



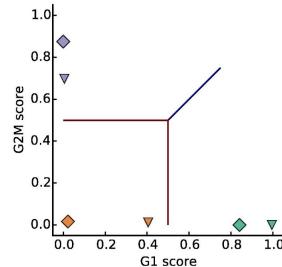
D LogReg



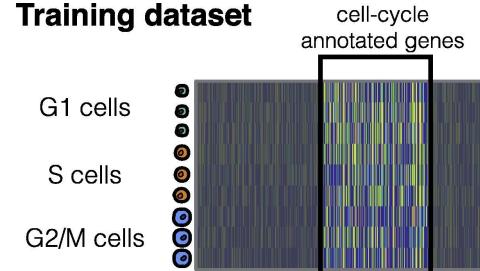
E SVM



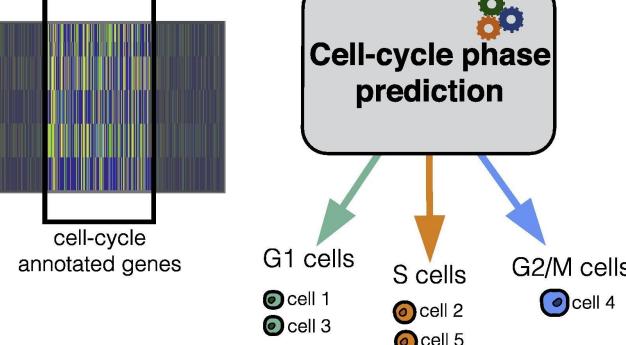
F Pairs



Training dataset

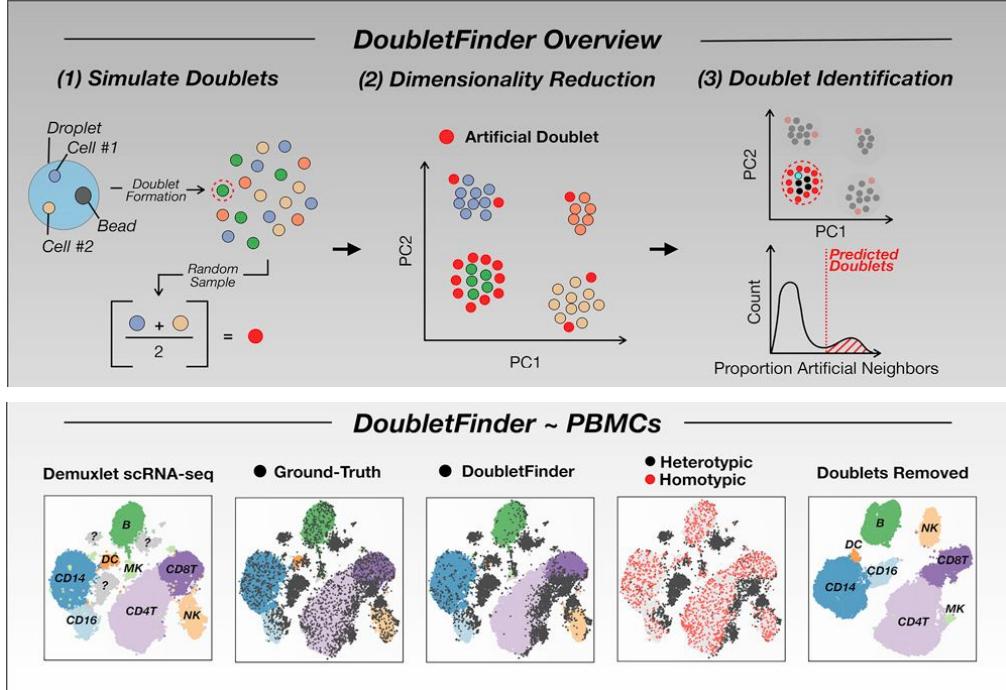


Testing dataset



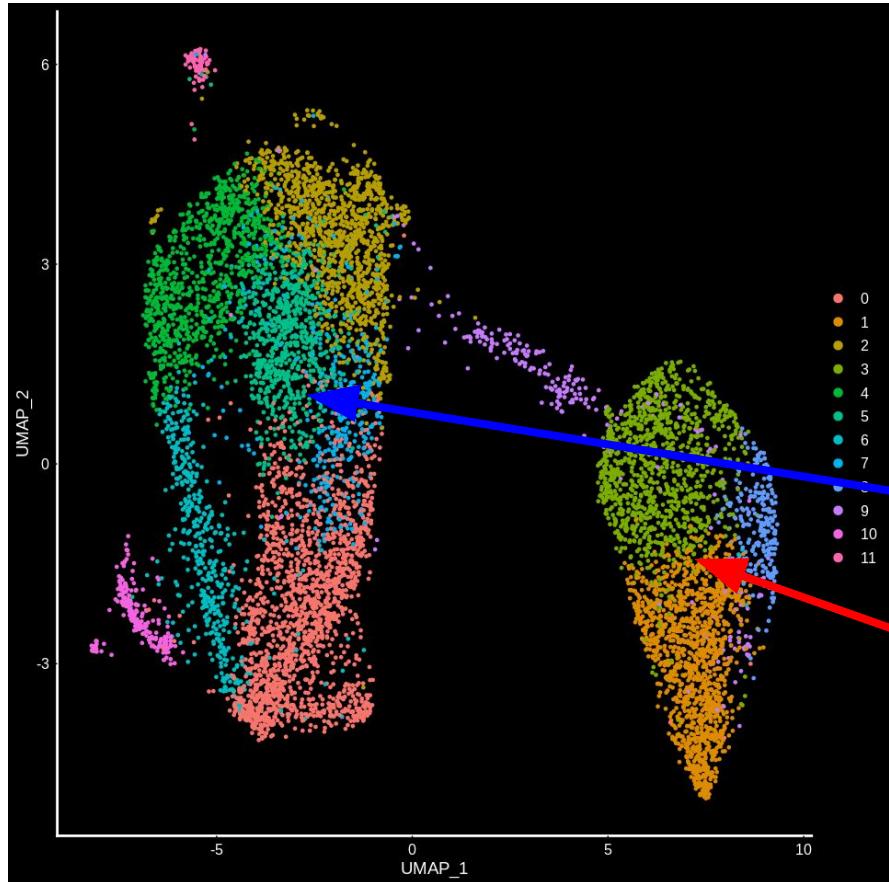
Other factors : Cell doublets

- 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



	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

Normalization : other biological factors

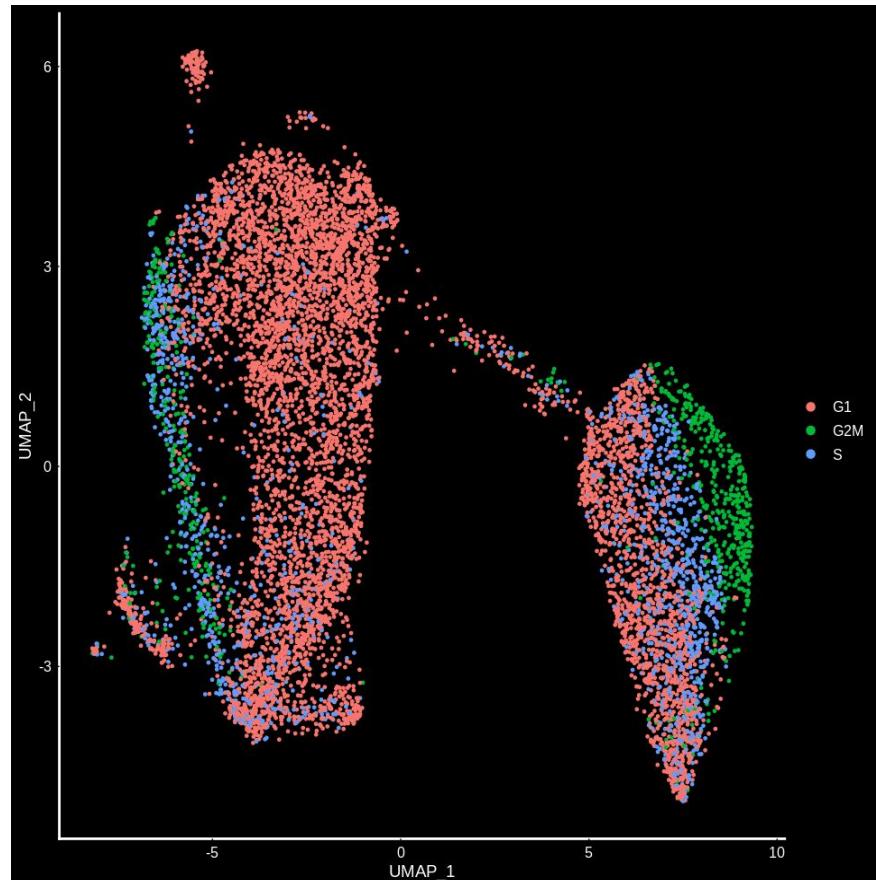
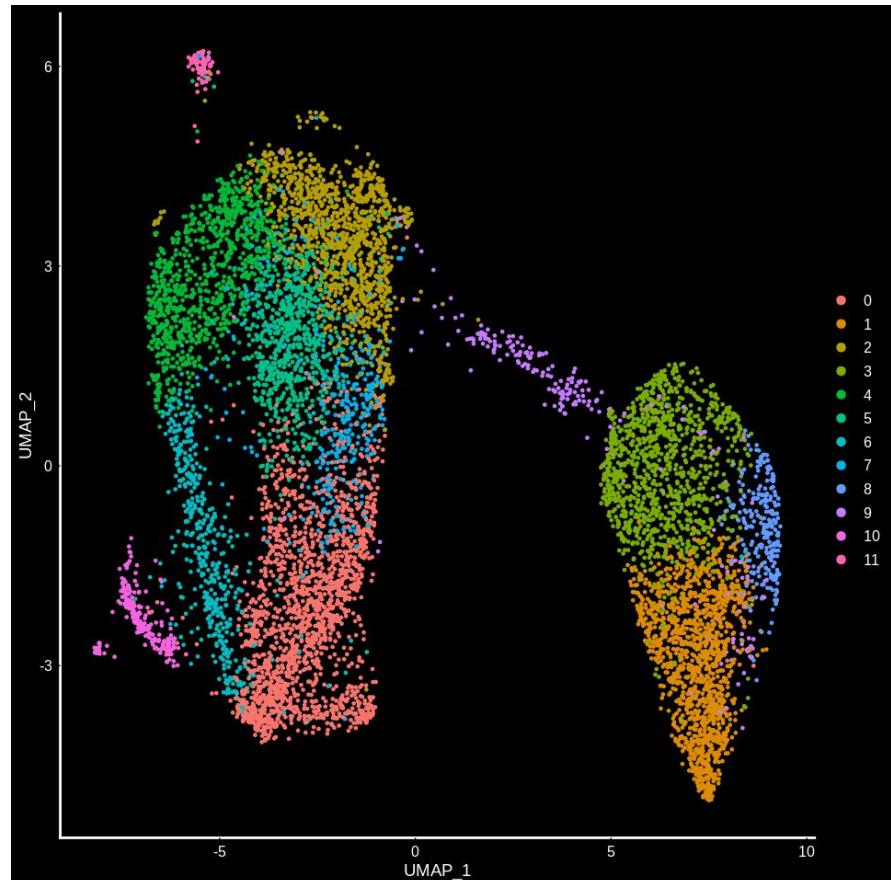


- 10X 3' scRNASeq
- Osteosarcoma metastasis
- 8911 cells x 18613 genes
- PCA (109 PCs retained)
- Louvain clustering
 - 12 clusters
- uMAP representation

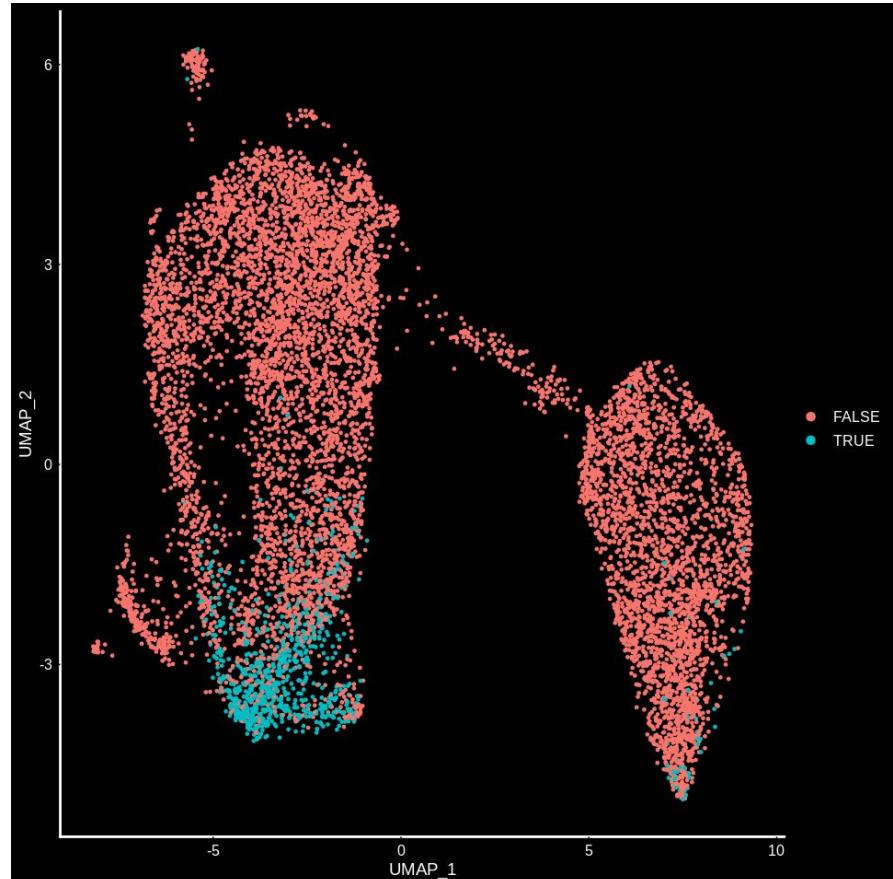
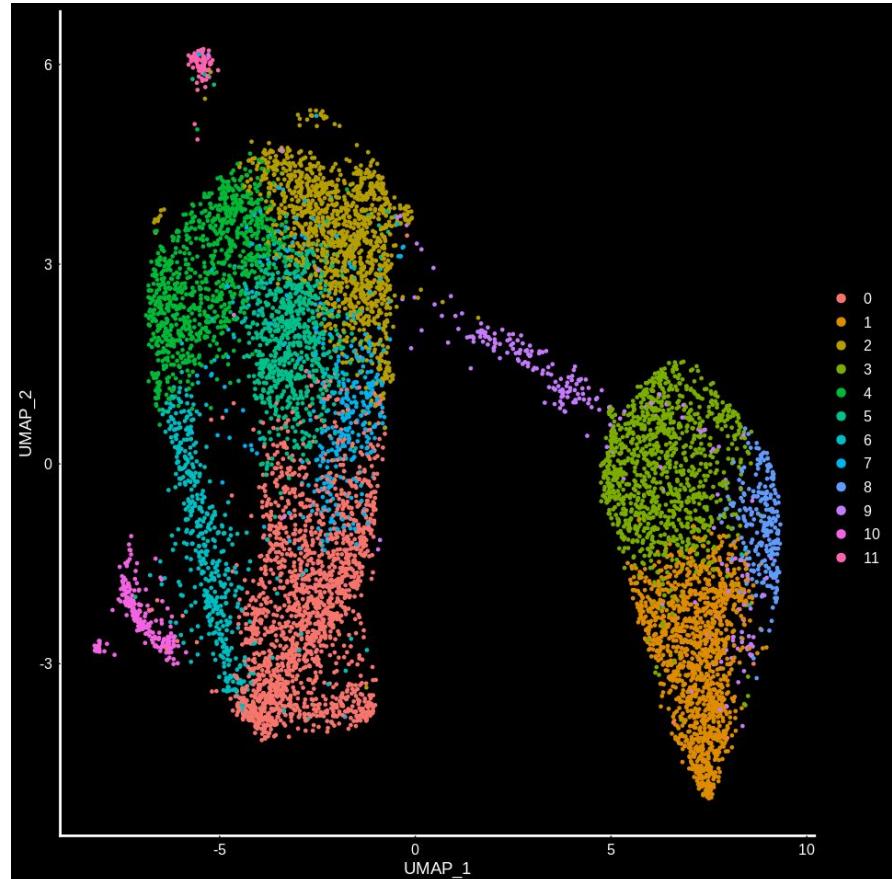
Osteoblasts

Osteoclasts

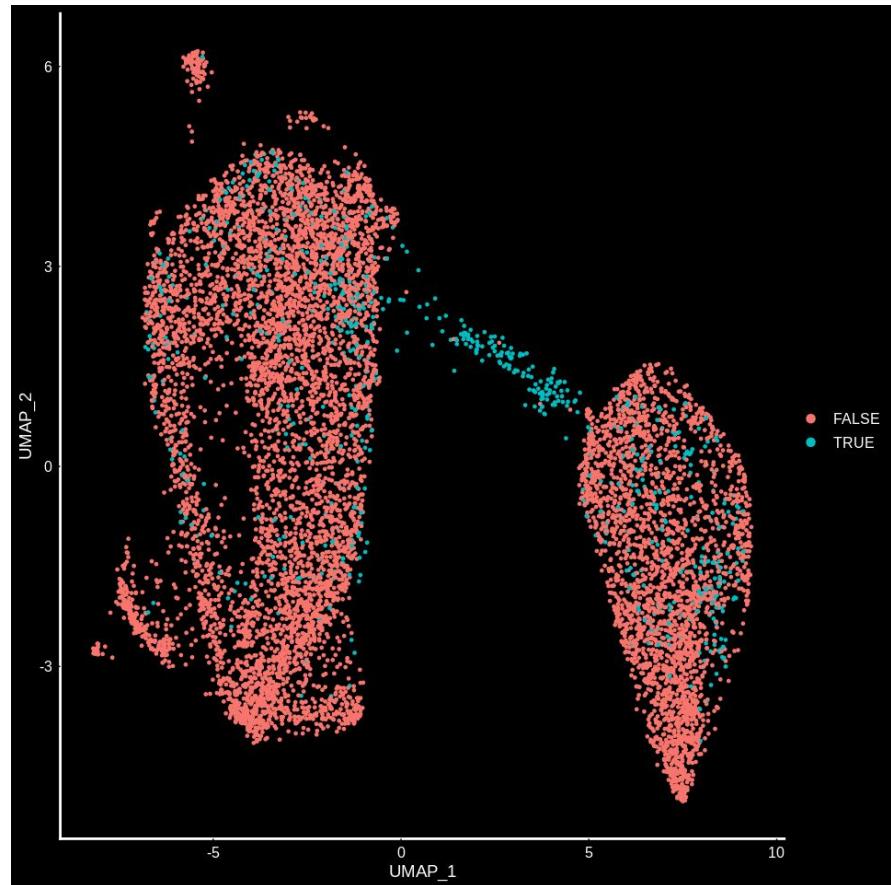
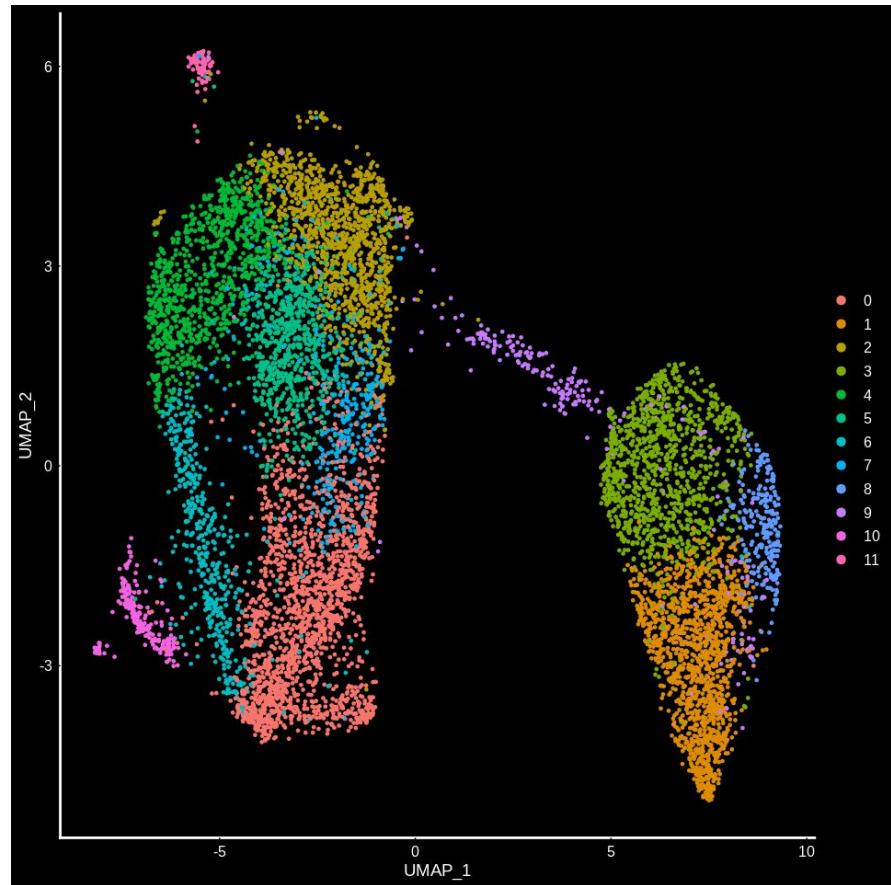
Bias : Cell cycle phases / scores



Bias : Dying cells status / score

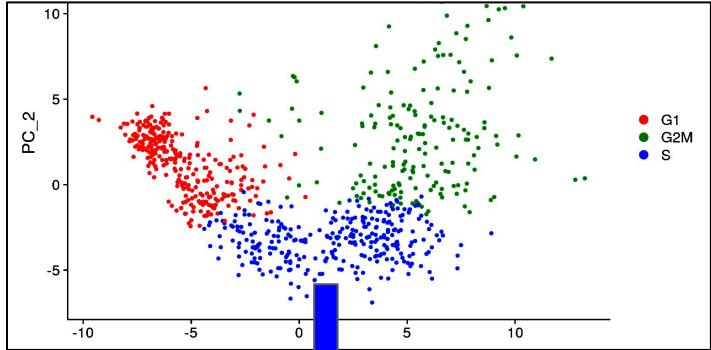


Bias : Cell doublet status / score

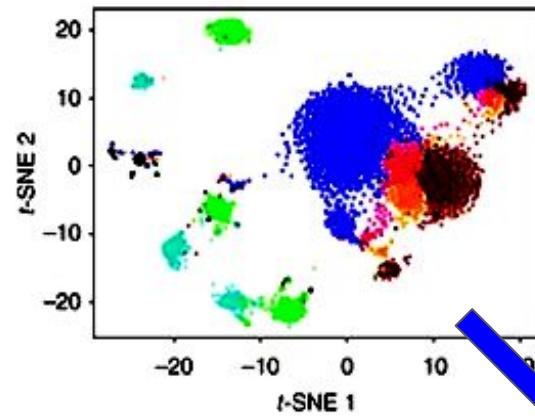


Bias normalization : regression / deblocking

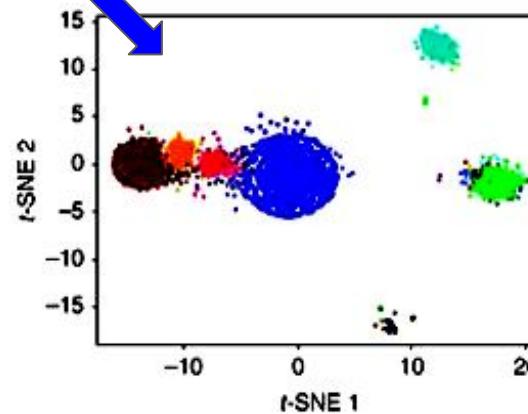
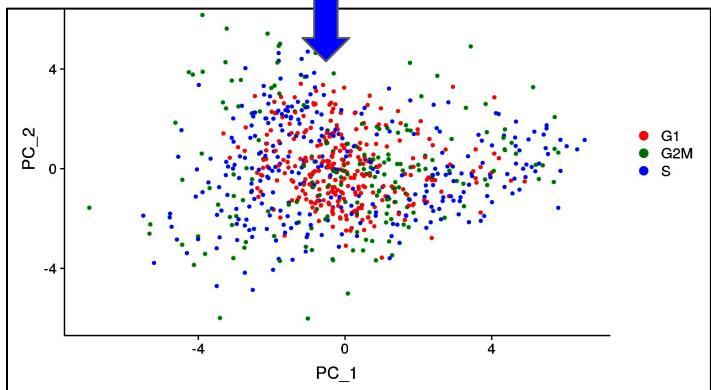
Ex : Cell cycle



Ex : Batch effect



Seurat tutorial



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)

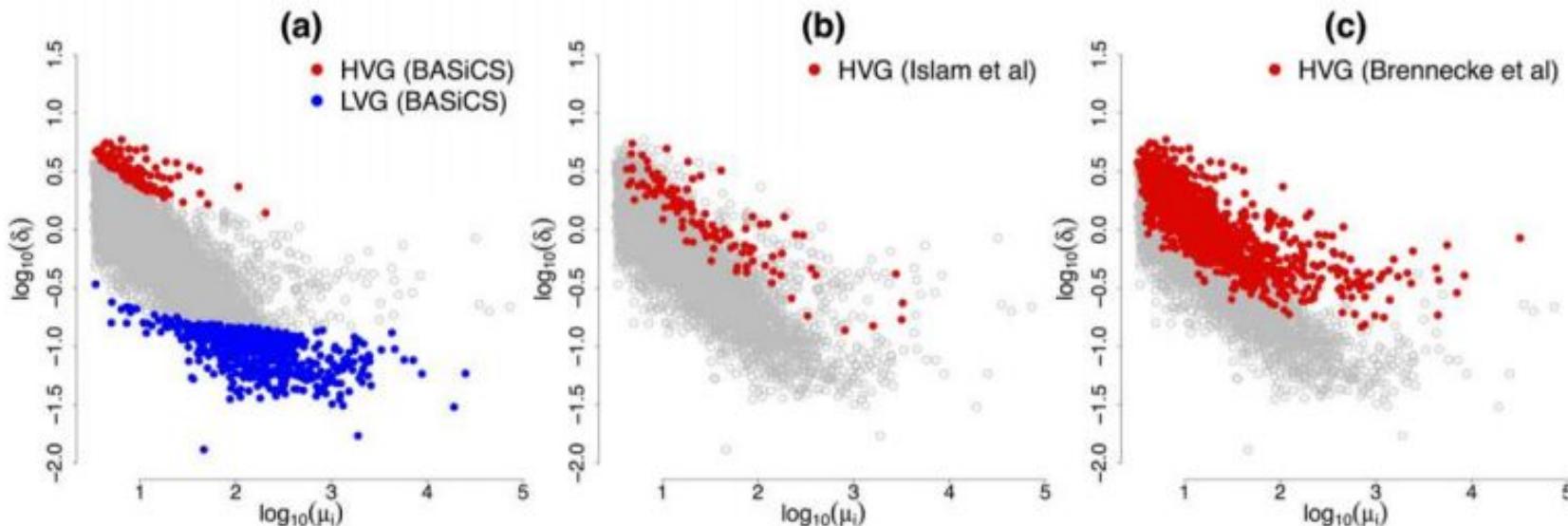


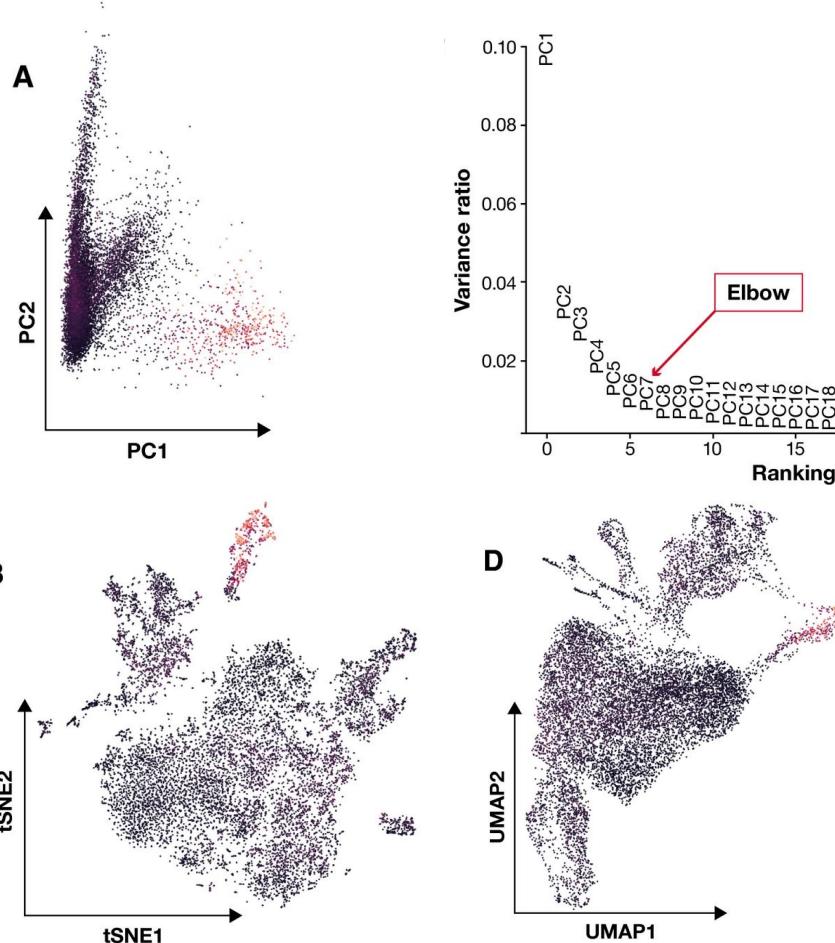
Fig 8. Comparison of HVG detection among different methods. For each of the 7,895 biological genes, posterior medians of biological cell-to-cell heterogeneity term δ_i (log scale) against posterior medians of expression level μ_i (log scale). While the methods described in [16] and [5] only provide a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

Dimensionality reduction : simplification

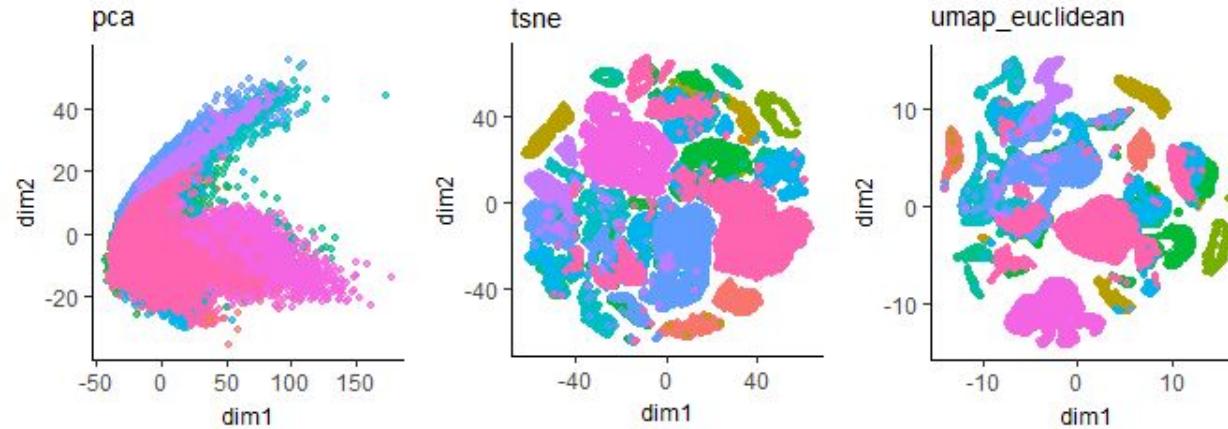
1. Need of an orthogonal space
2. Minimize curse of dimensionality
3. Filter out noise
4. Allow visualization
5. Reduce computational load

Popular methods used for single-cell data analysis:

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



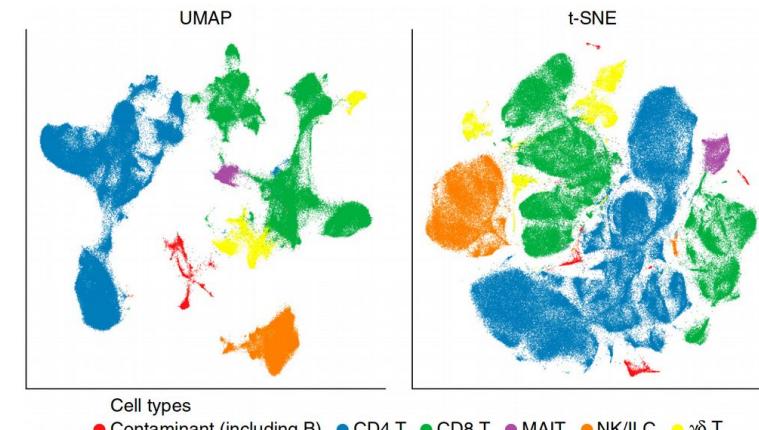
Dimensionality reduction : PCA / tSNE / uMAP ?



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

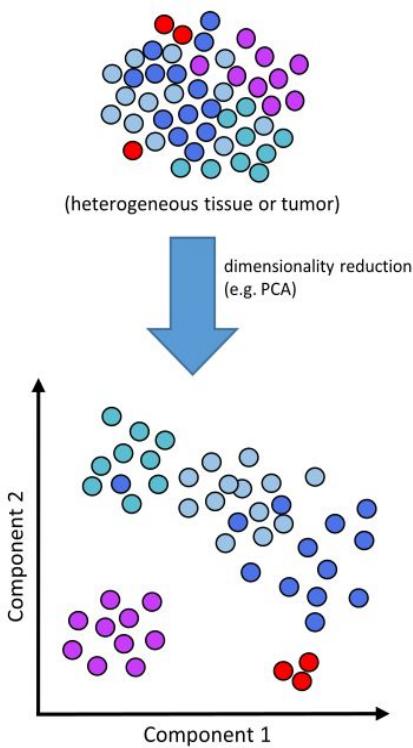
uMAP > tSNE :

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

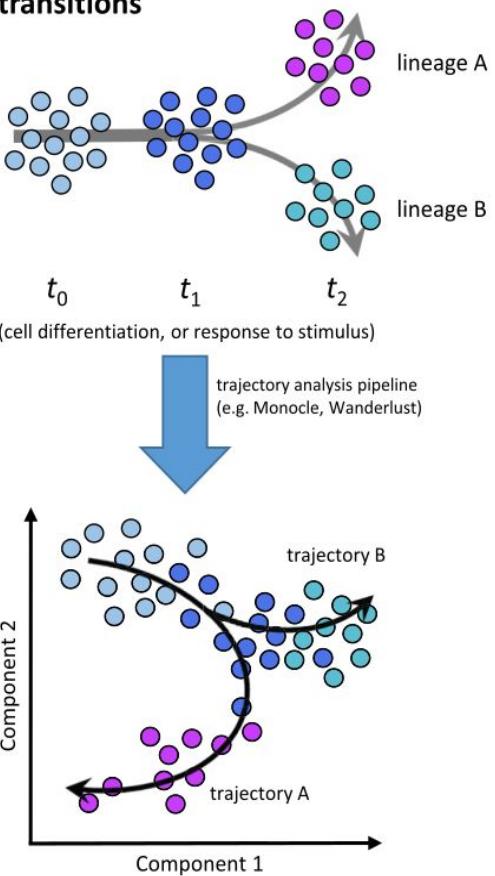


From a reduced space to ...
... actually what you initially wanted !

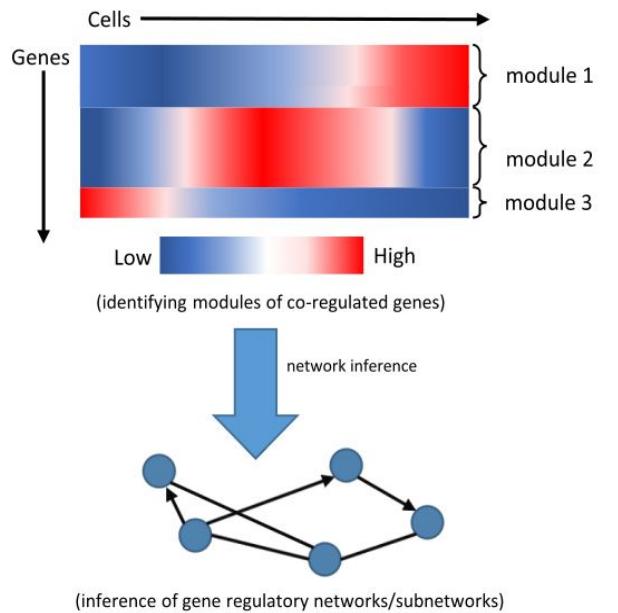
a) Deconvolving heterogeneous cell populations



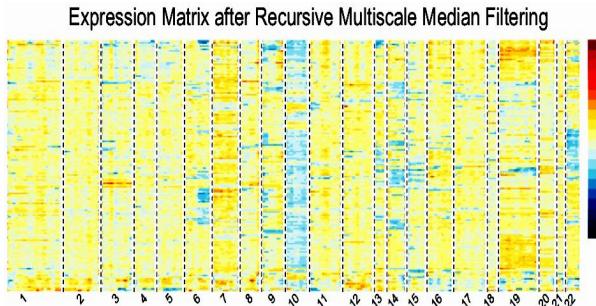
b) Trajectory analysis of cell state transitions



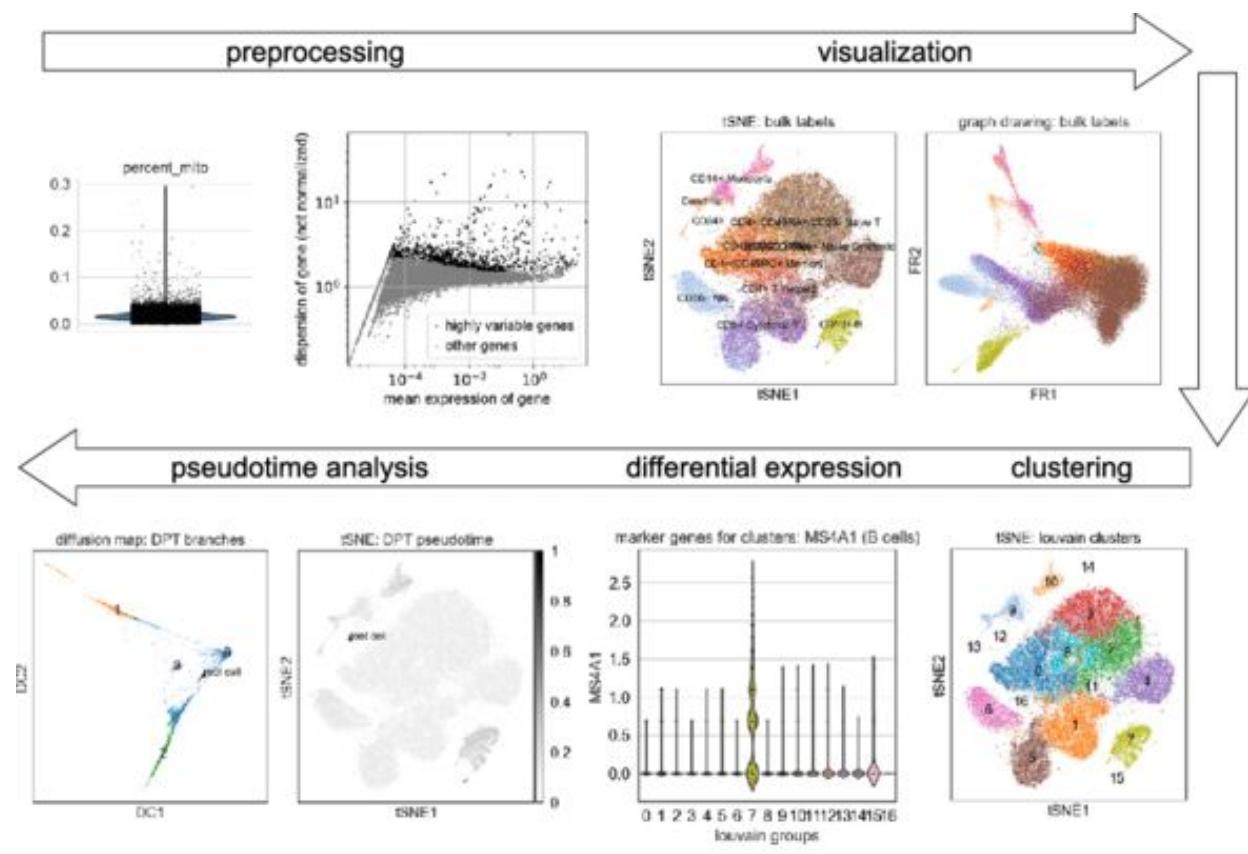
d) Network inference



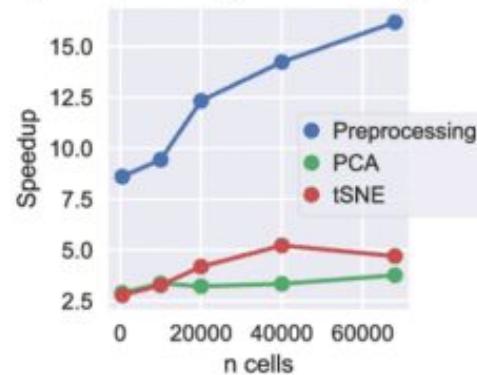
e) Copy number estimation



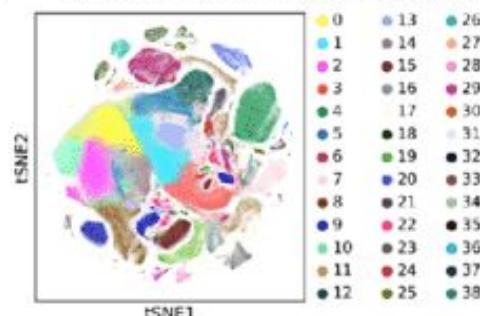
The all-in-one Python toolbox : Scanpy



b Speedup: Scanpy vs. Cell Ranger R



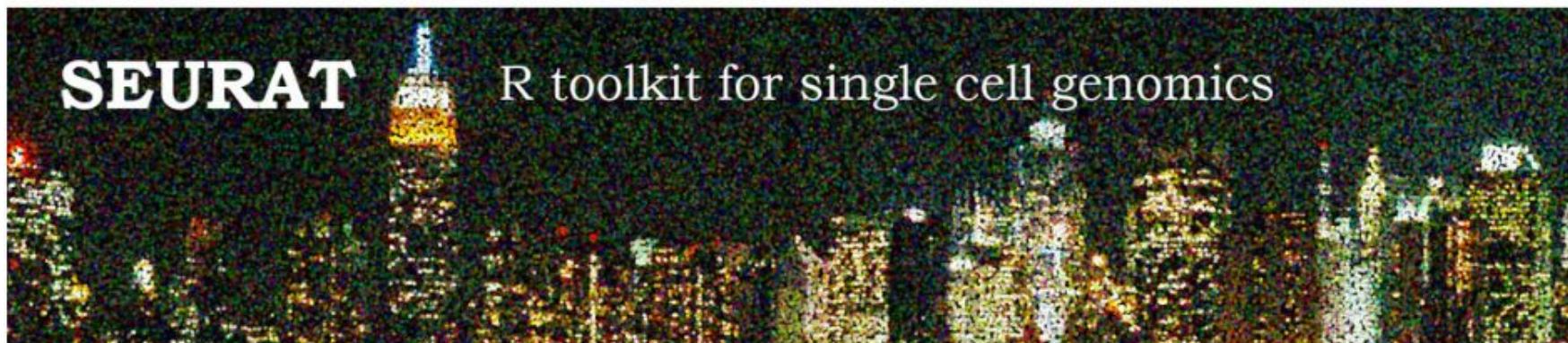
c tSNE of clustered 1.3 million cells



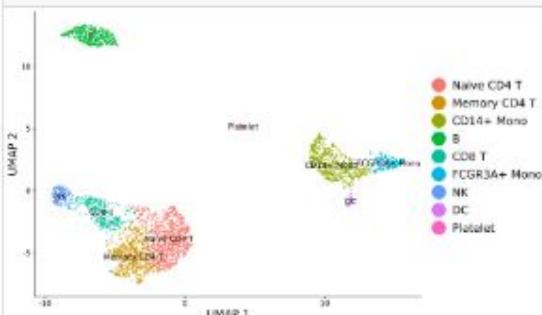
The all-in-one R toolbox : Seurat (v3)

HOME NEWS PEOPLE RESEARCH PUBLICATIONS SEURAT JOIN/CONTACT

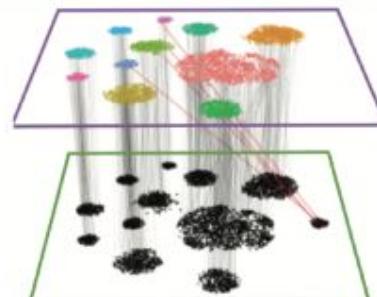
SINGLE CELL
GENOMICS DAY



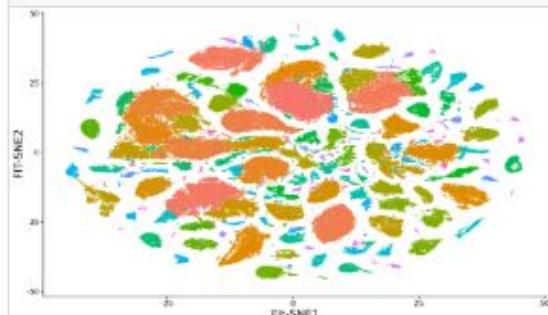
Guided tutorial --- 2,700 PBMCs



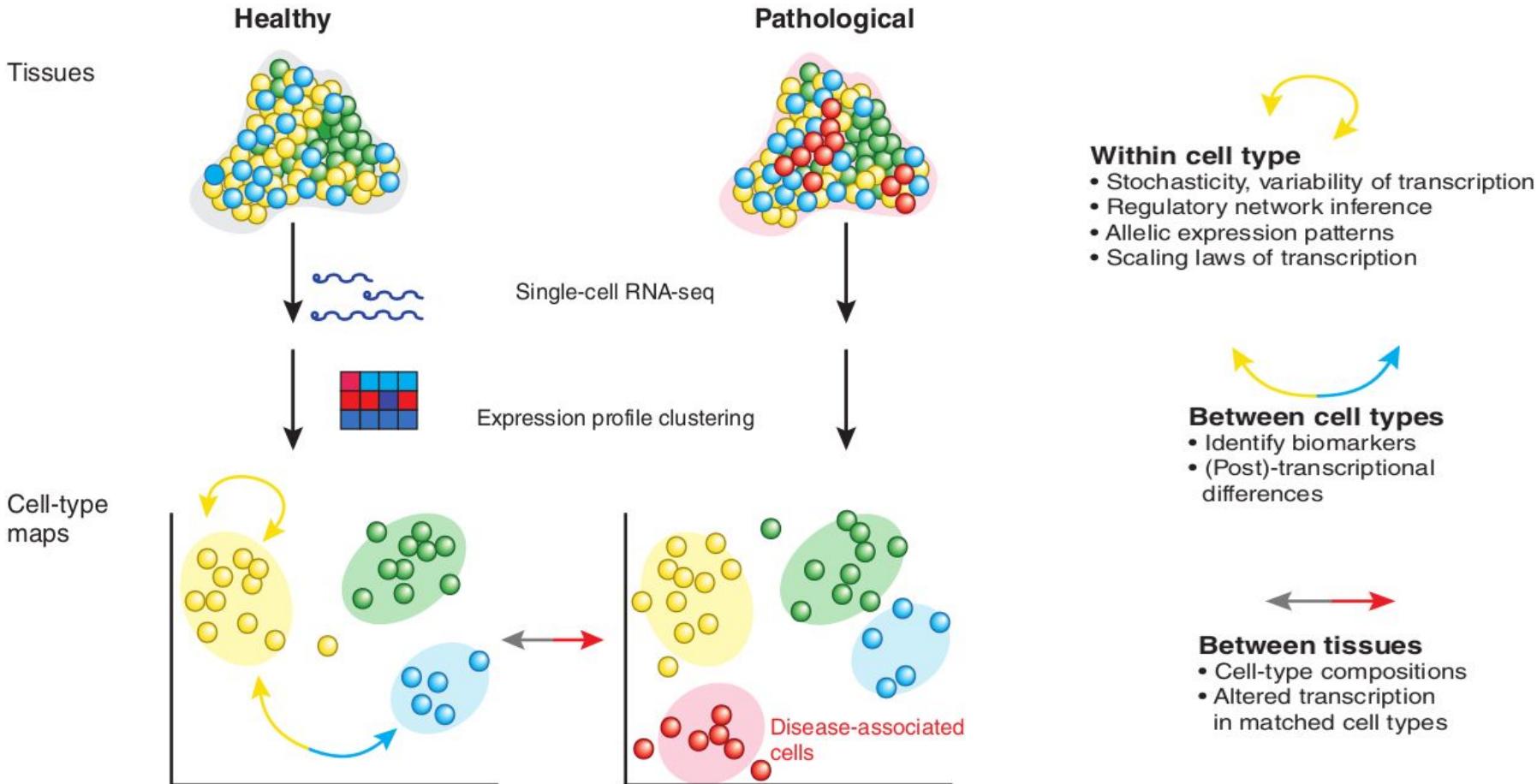
Multiple Dataset Integration and Label Transfer



Mouse Cell Atlas, 250K cells

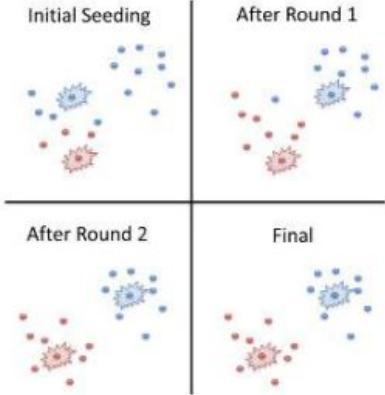


Cell clustering

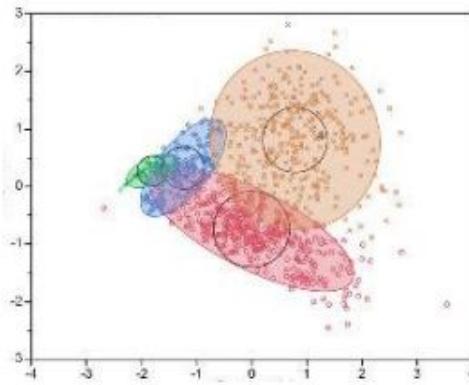


Cell clustering : methods

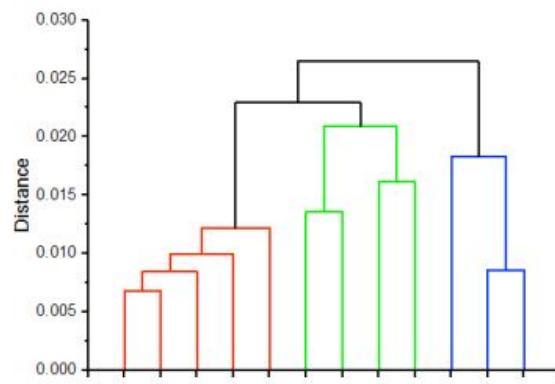
1) K-means based



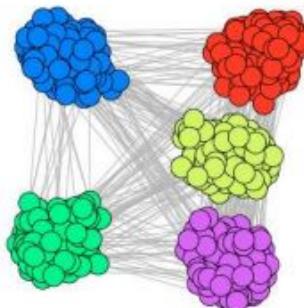
3) Model-based clustering (Mclust)



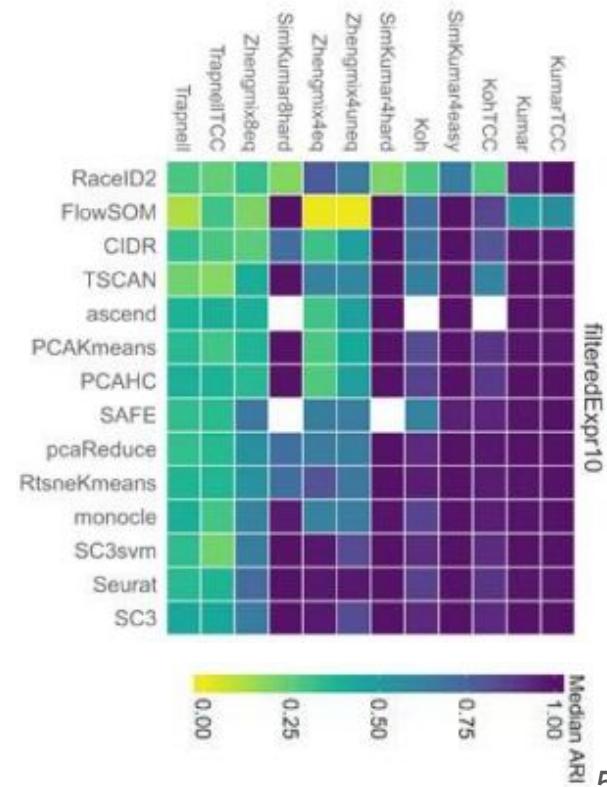
2) Hierarchical clustering



4) Graph-based clustering

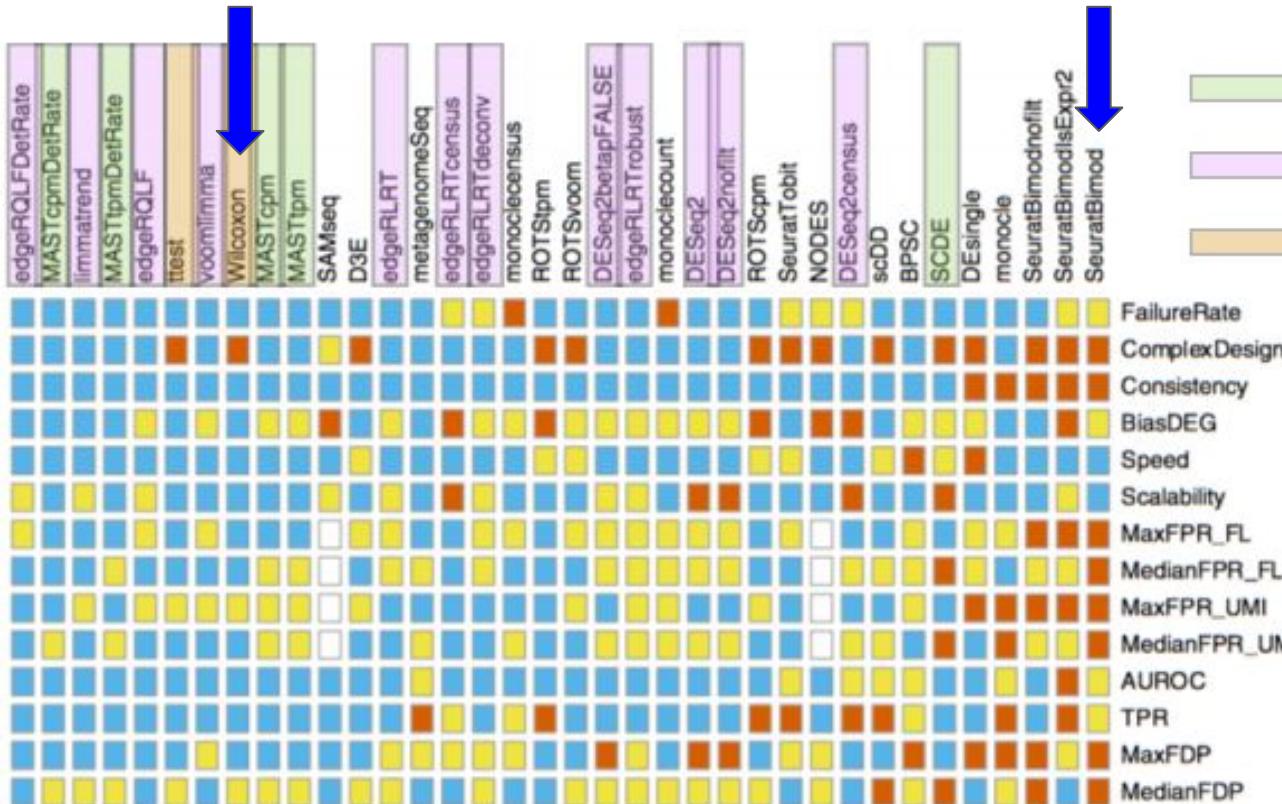


5) Single-cell specific methods



Differential expression analysis

Seurat v3



Seurat v2

Bayesian 3-component model adapted to single-cell data

Methods borrowed from bulk-RNA-seq

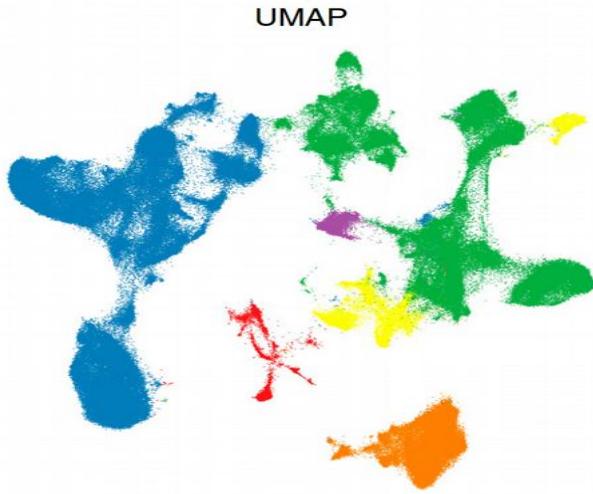
Naïve approaches

Good

Intermediate

Poor

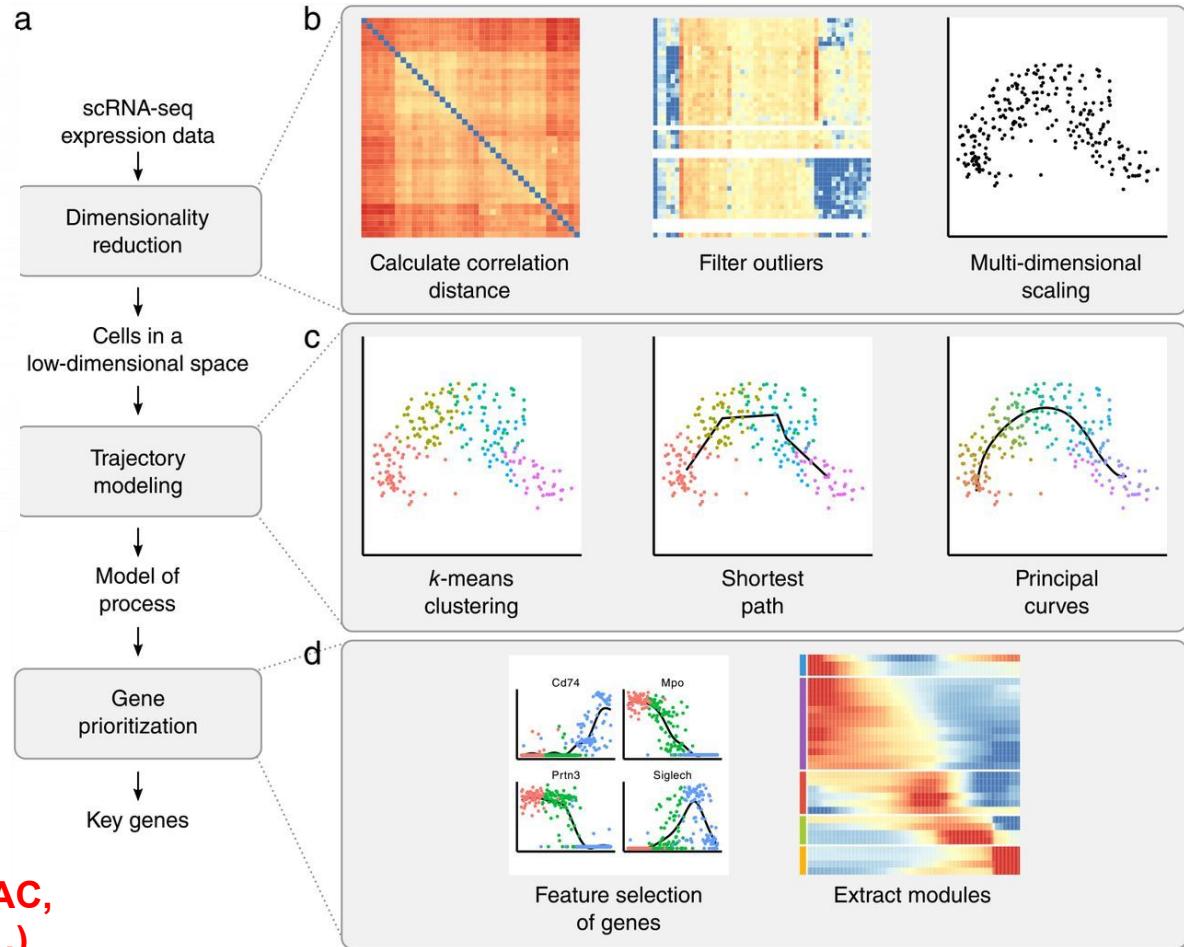
Cell trajectory : methods



Most adopted tools :

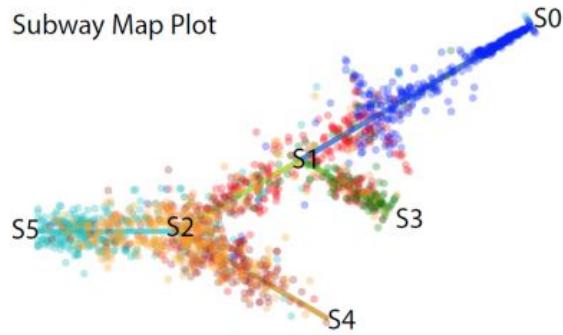
- Monocle 3
- PAGA
- STREAM
- Scorpius
- Slingshot

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

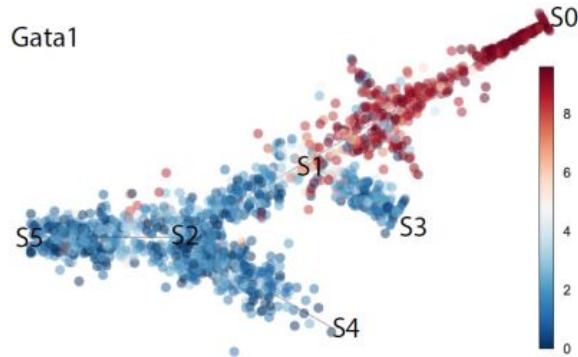


Cell trajectory : visualization

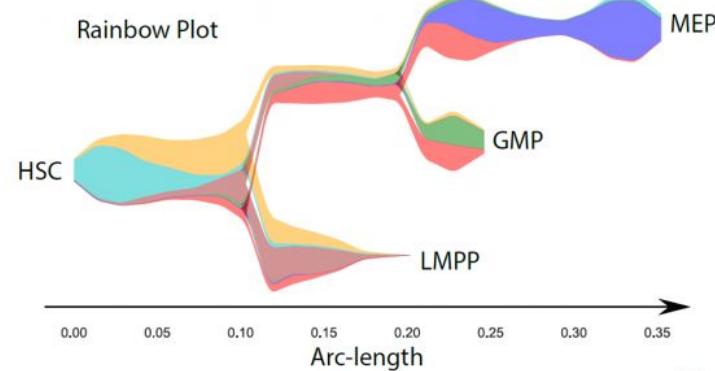
Cell distance to path + cell types



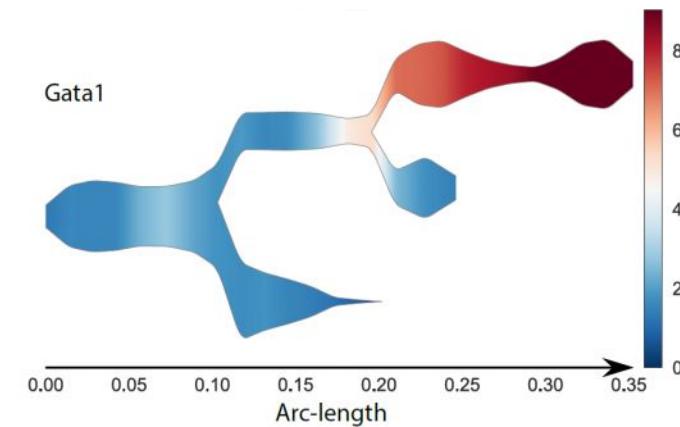
Cell distance to path + gene expression



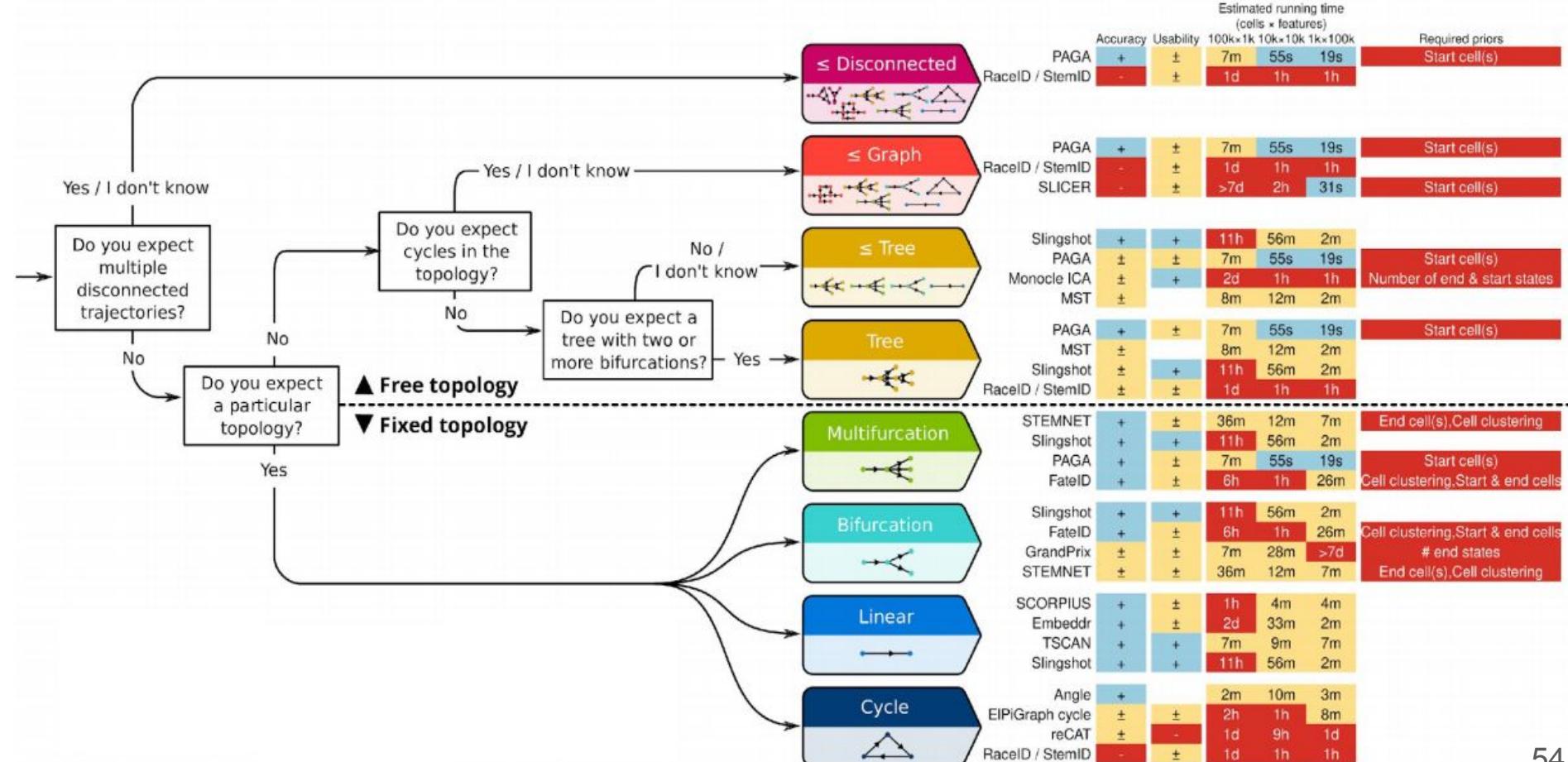
Cell types density



Cells density + gene expression

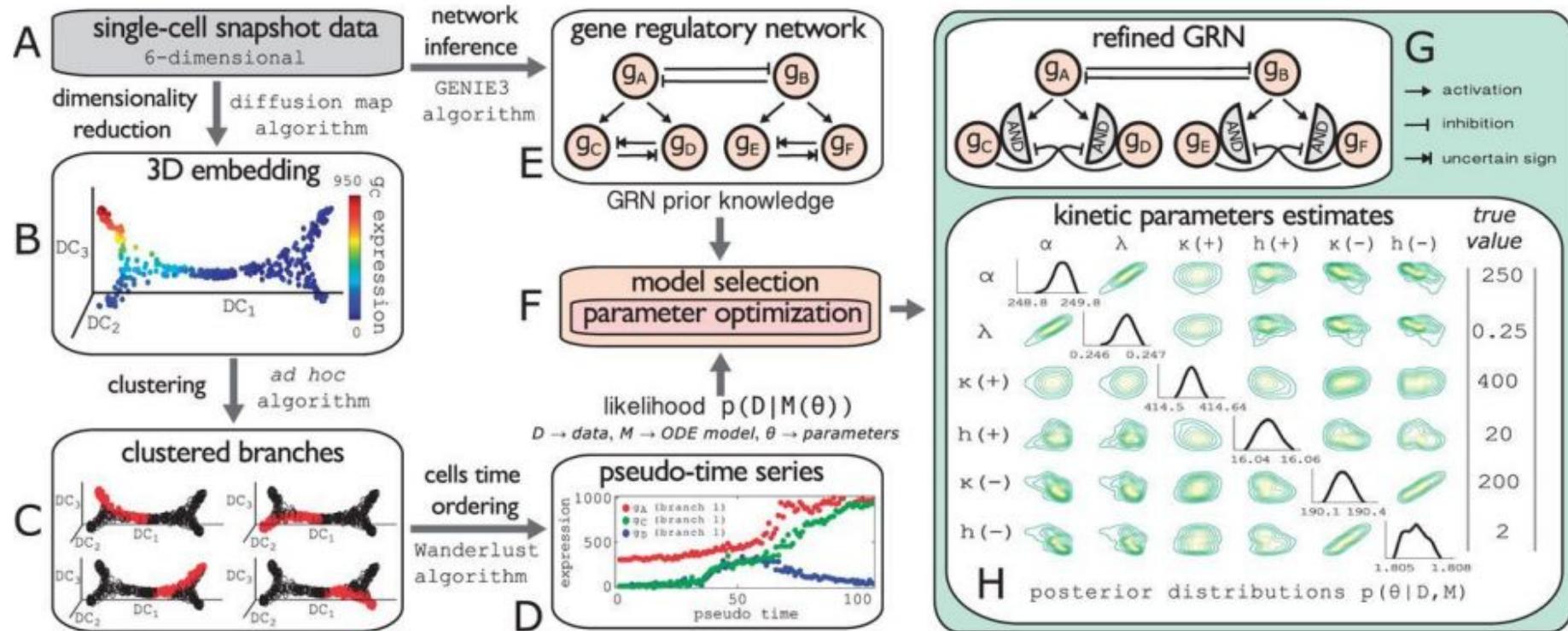


Cell trajectory : Multiple contexts



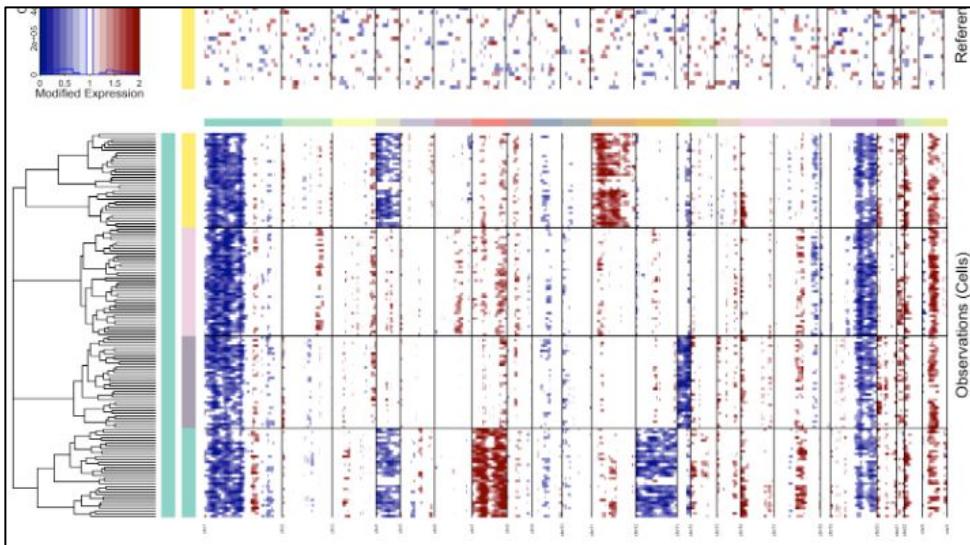
Network inference

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



Copy number estimation from scRNASeq

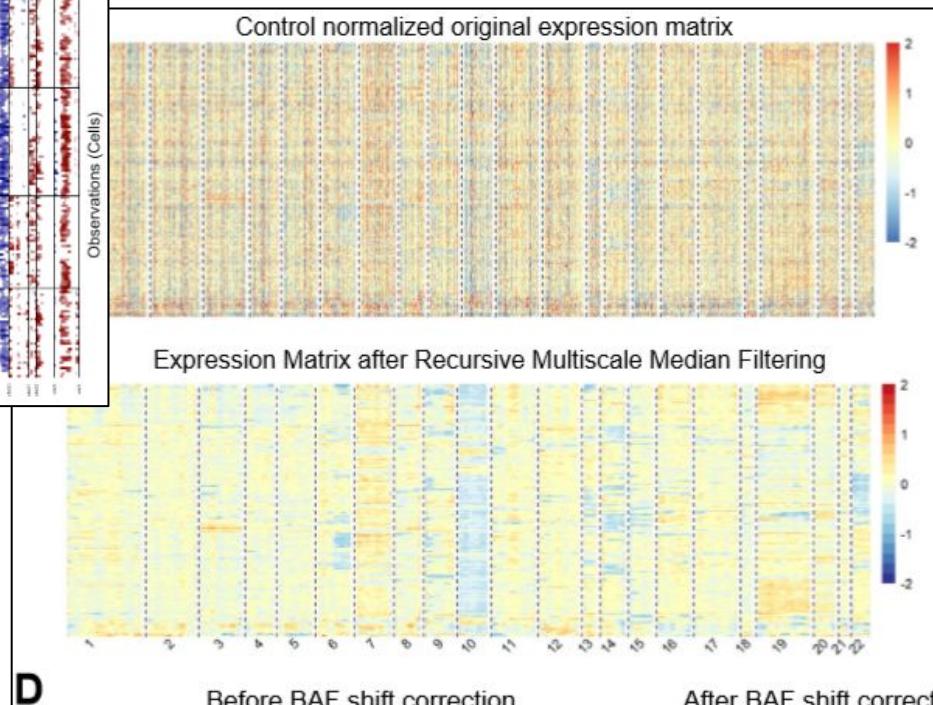
InferCNV (Broad Institute)



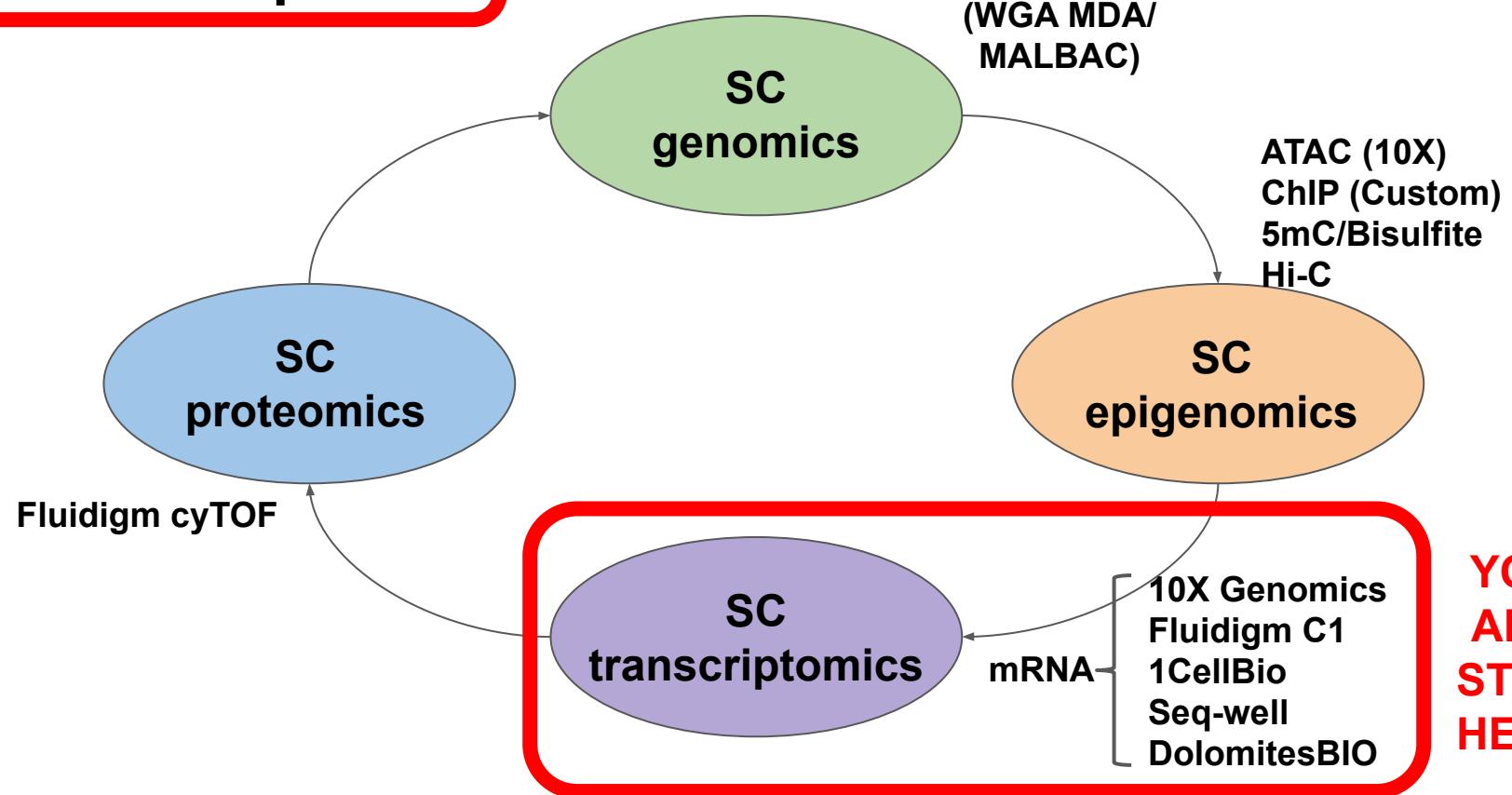
WARNING :

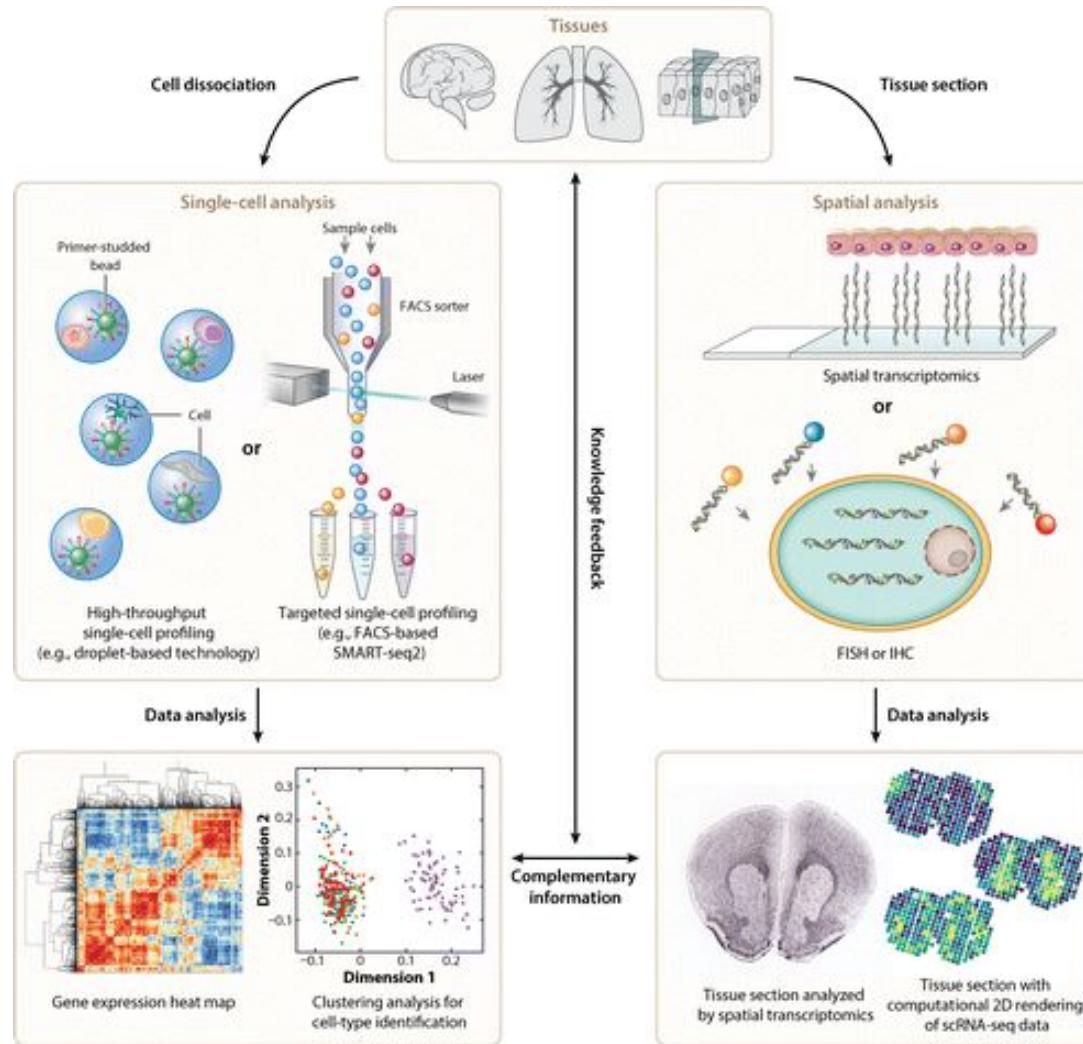
- Coarse grain (> 10 Mb)
- Requires > 75,000 reads / cell

CaSpER (Armanci et al, BioRxiv 2019)

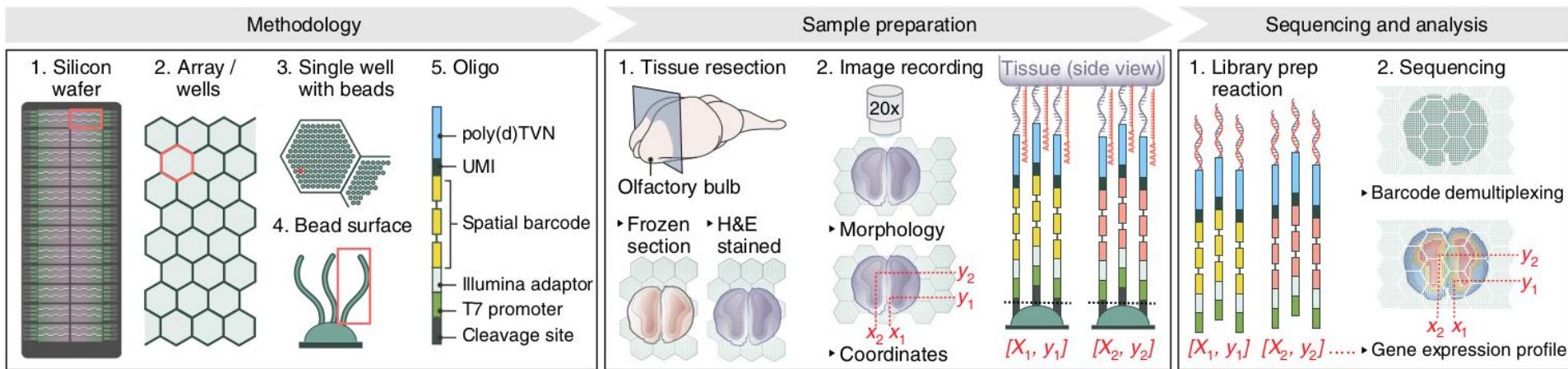


Spatial Single Cell RNAseq



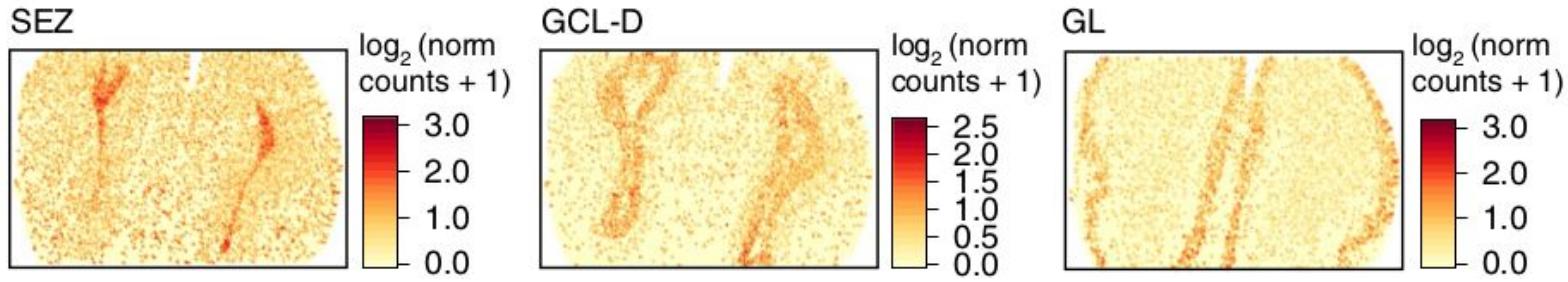


Illumina “High Definition Spatial Transcriptomics”

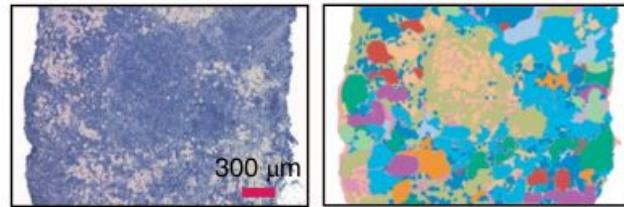


- 2,893,865 individual barcoded beads
- 1.4 M wells
- Well diameter $\sim 2 \mu\text{m}$
 - << median cell diameter (20 μm)
 - $\sim 1,400 \times$ higher resolution than “standard” ST
 - $\sim 25 \times$ compared to SLIDE-seq
- Array reading time $\sim 3 \text{ H}$
- Challenging analysis strategy (low capture rate) ...
- Commercially available in 2020

Illumina HDST



H&E Annotations



- Fatty tissue, immune/lymphoid
- Fibrous tissue, invasive cancer
- Fibrous tissue, immune/lymphoid
- Invasive cancer, immune/lymphoid
- Immune/lymphoid
- Fatty tissue, fibrous tissue, invasive cancer
- Fibrous tissue
- Fibrous tissue, invasive cancer, immune/lymphoid
- Fatty tissue
- Fatty tissue, fibrous tissue, invasive cancer, immune/lymphoid
- Invasive cancer
- Fatty tissue, invasive cancer, immune/lymphoid

C

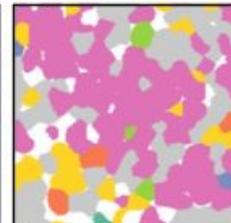
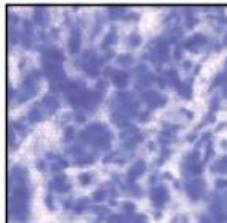
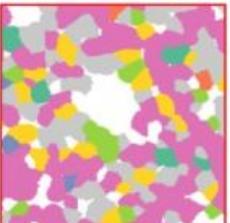
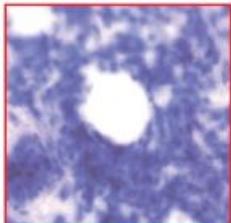
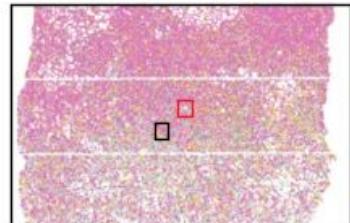
Cell types
in sn-like data

H&E
enlargement

sn-like
enlargement

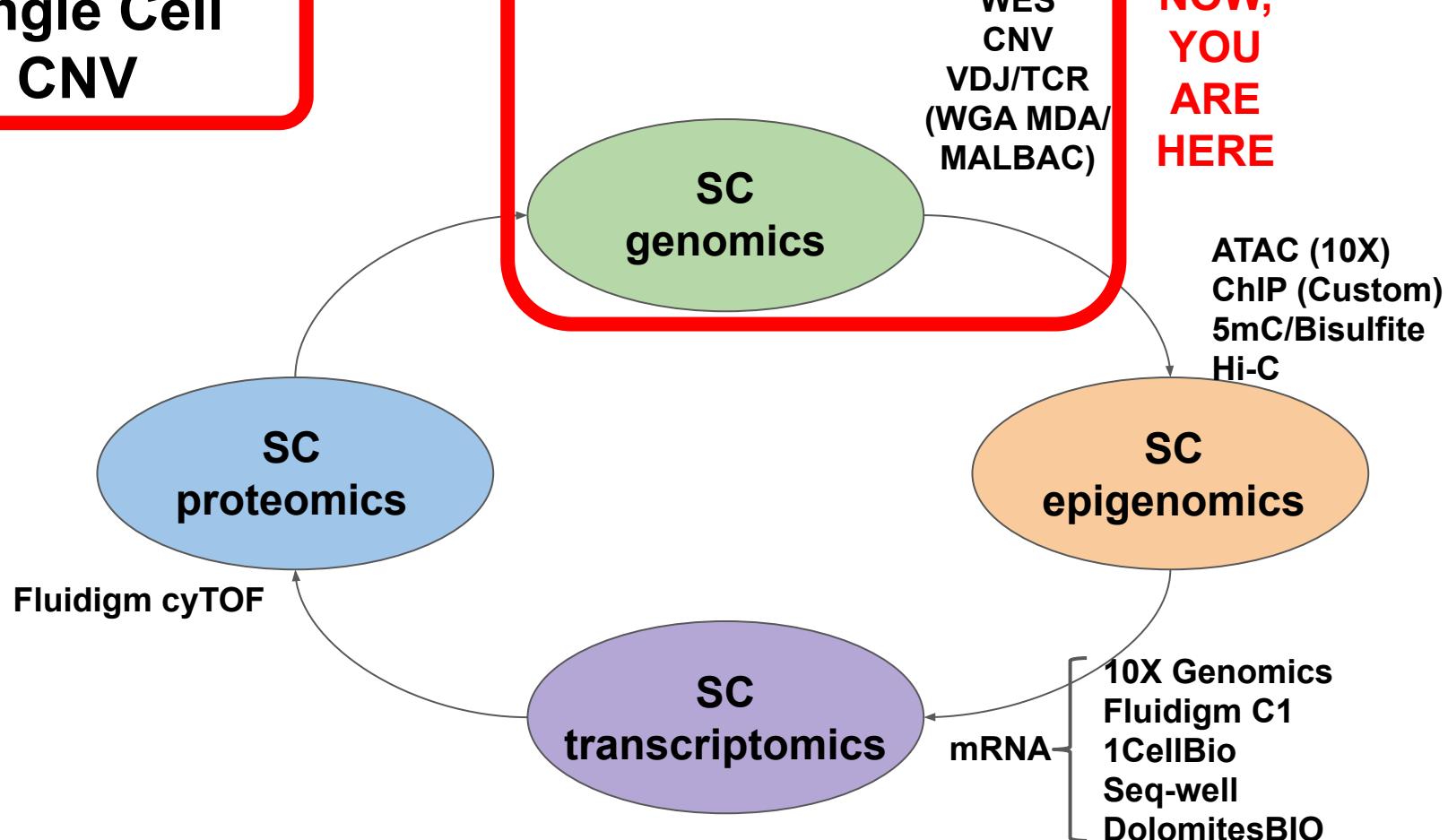
H&E
enlargement

sn-like
enlargement



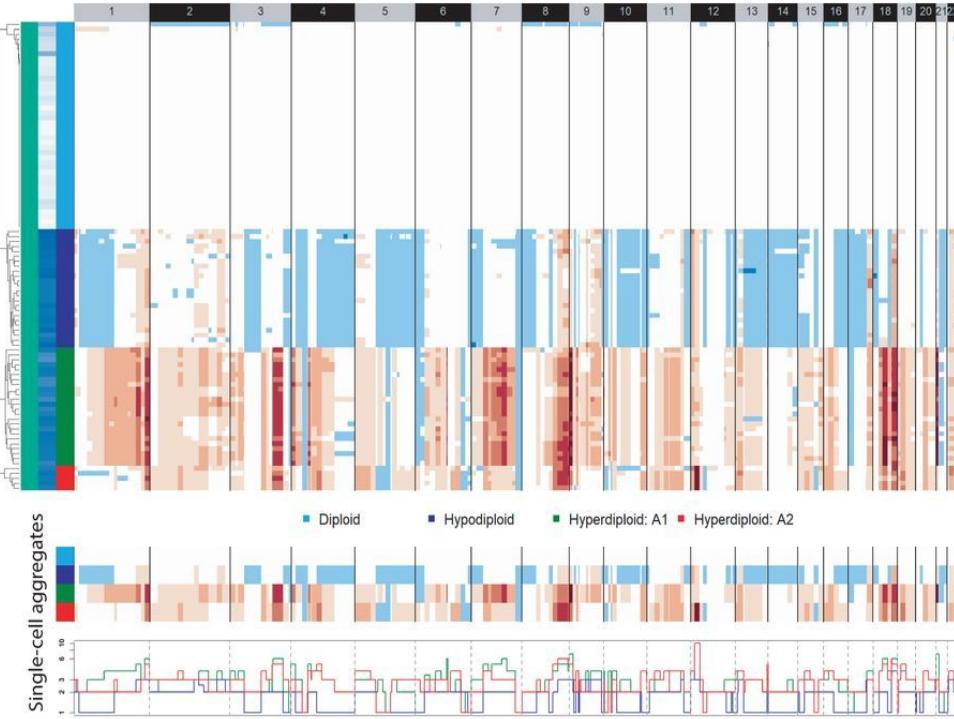
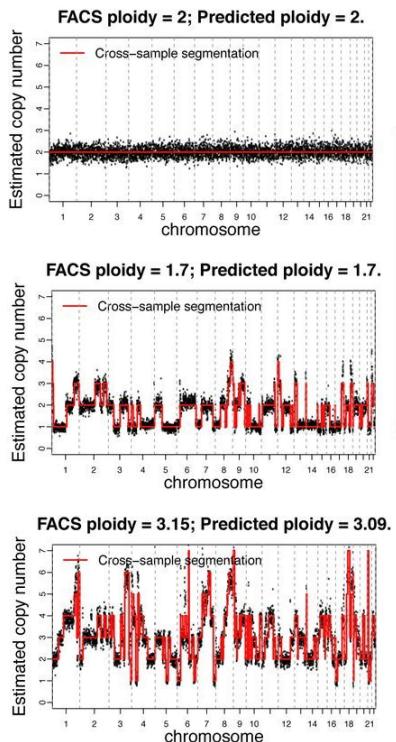
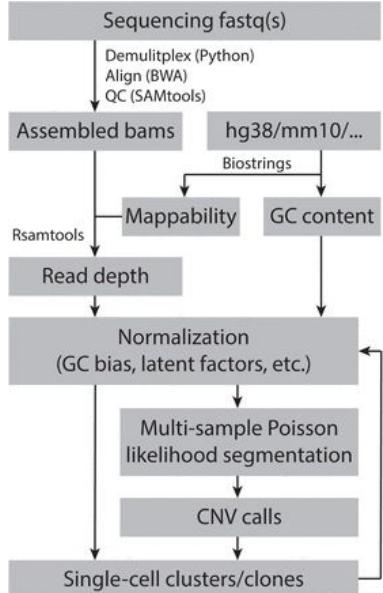
- T cells
- B cells
- Endothelial cells
- Epithelial cells
- Macrophages
- Stroma
- Unassigned nucleus

Single Cell CNV



NOW,
YOU
ARE
HERE

scCNV results (SCOPE)

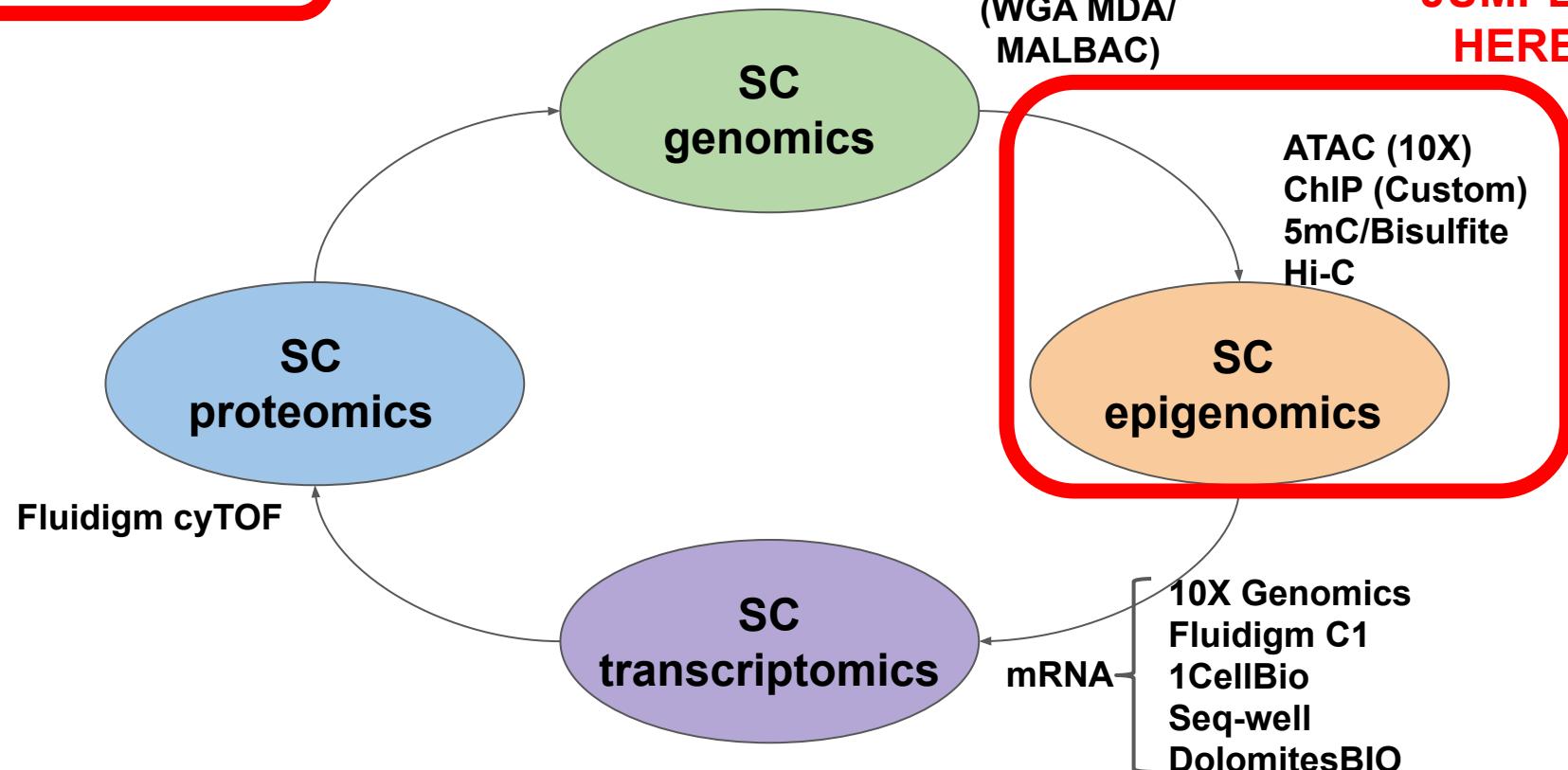


WARNING :

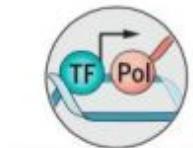
- Limited resolution : > 2 Mb (binning)
- Requires > 750,000 reads / cell

Single Cell Epigenomics

YOU
JUMPED
HERE

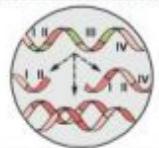


Overview of scEpigenomics techniques

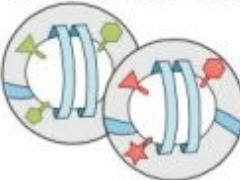


Transcription factor binding

TF binding interacts with DNA methylation and chromatin accessibility



Transcription and RNA maturation



Histone modifications

Modifications can be active marks (e.g., H3K4me3 in green) or repressive marks (e.g., H2K27m3 in red)



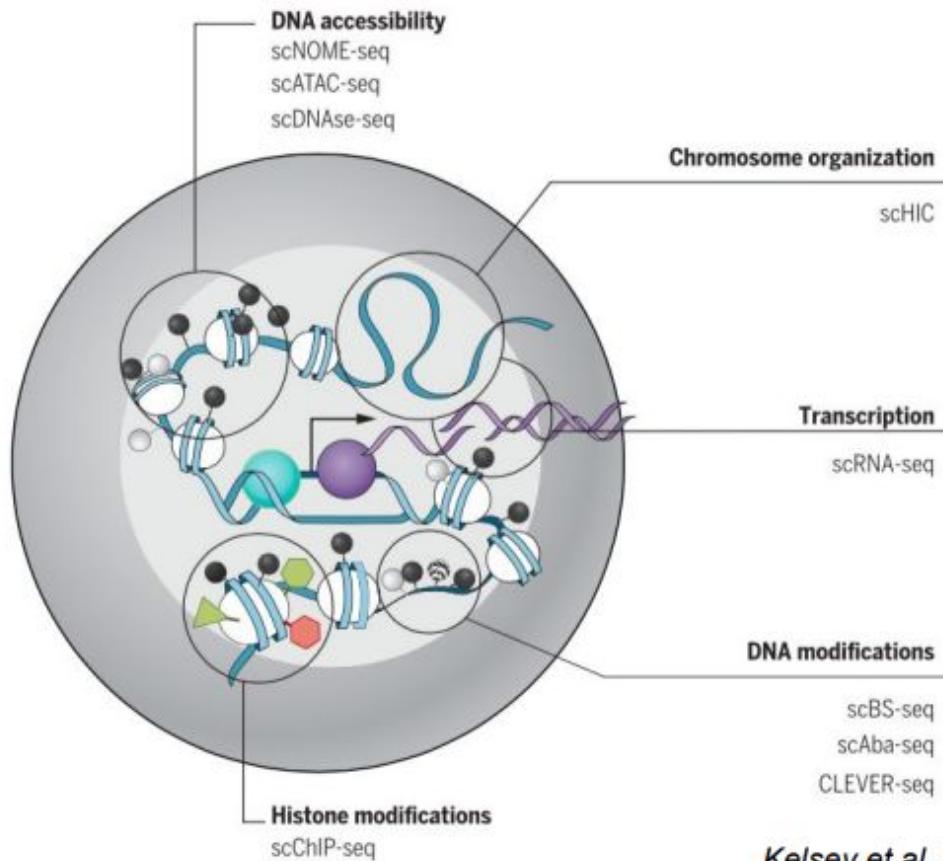
DNA modifications

● C ● 5mC
● 5hmC / 5fC / 5caC



Chromosome organization

Higher-order chromatin organization into LADs and TADs



Single Cell (RNAseq) Resources

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

Mapping the Mouse Cell Atlas by Microwell-Seq

Graphical Abstract

The graphical abstract illustrates the workflow for constructing the Mouse Cell Atlas. It starts with a green silhouette of a mouse labeled '>>40 Mouse organs and tissues'. An arrow points from the mouse to a cluster of colored shapes representing 'The Mouse Cell Atlas'. Another arrow points from the mouse silhouette to a purple circle containing the text '>>400,000 Single Cell mRNA-seq'. From this circle, an arrow points to a vertical stack of blue rectangles labeled 'Microwell-seq' with the sub-label 'Cells'. Below this stack is a grey rectangle labeled 'Wash Out'. A second stack of blue rectangles labeled 'Beads + Cells' is shown below the first. From the 'Beads + Cells' stack, an arrow points to a purple circle containing the text 'Single Cell DGE Data'. From this circle, an arrow points to a grey cylinder labeled 'scMCA' with the sub-label 'Cell Type Identification'.

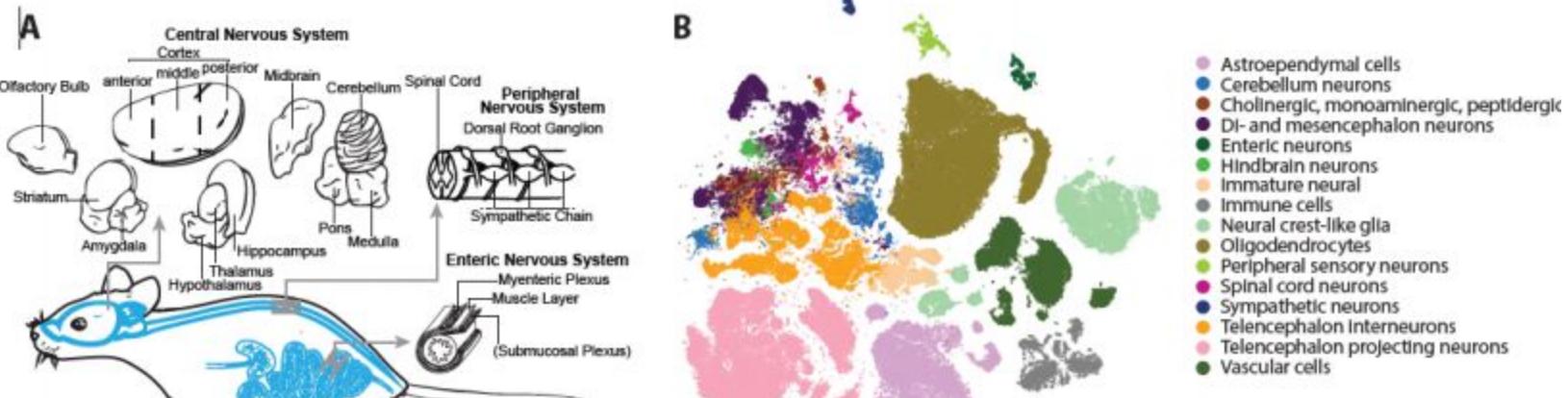
Resource

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

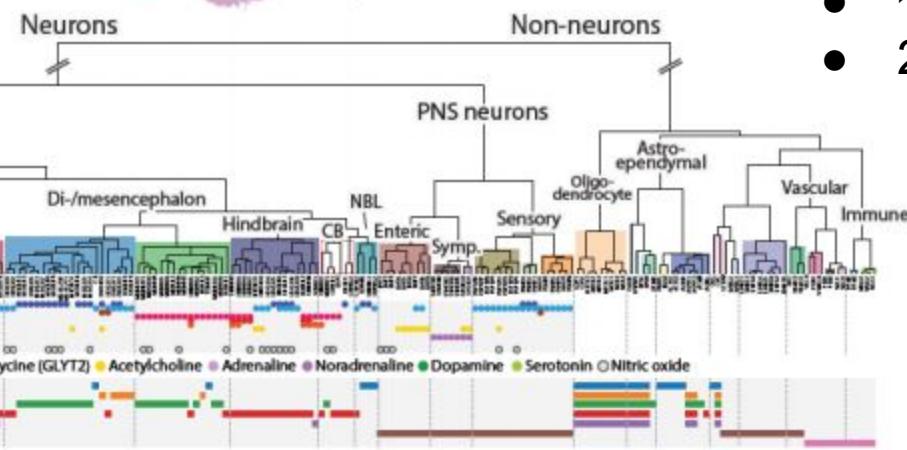
Correspondence
xhan@zju.edu.cn (X.H.),
ggj@zju.edu.cn (G.G.)

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)



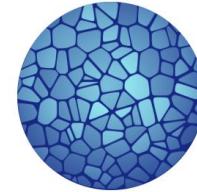
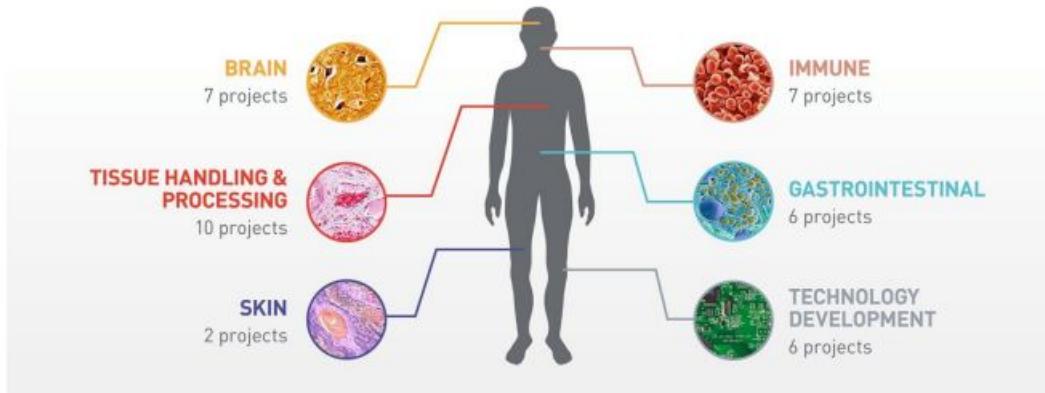
C



The Human Cell Atlas (humancellatlas.org)

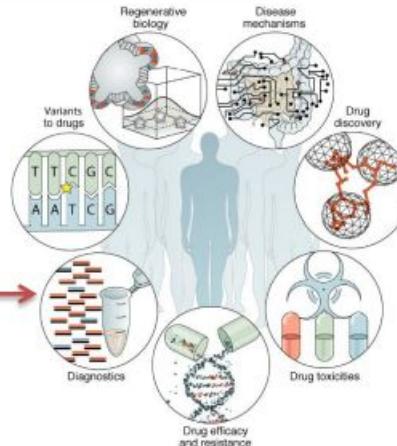
MAPPING THE BASIC UNITS OF LIFE

CZI proudly supports 38 new projects in these six areas for the Human Cell Atlas.



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



SCHNAPPS : A R-shiny app for biologists



SCHNAPPS

Information:

- Clustering: Clustering was performed with t.SNE followed by identification using DBSCAN
- Cluster 0: Cells that cannot be assigned to any cluster
- 3D Plot: Enter gene name to visualize expression in a single cell
- 2D Plot: Pick a cluster, highlight cells of interest to download gene expression matrix

Enter gene CD52

comma separated list of genes for UmiCountPerGenes comma separated list of genes for UmiCountPerGenes2

X tsne1 Y tsne2 color Gene.count

l-Cluster1 CD52-Cluster2 CD52-Cluster3 CD52-C

tsne2

tsne1

show more options

CD52

Expression

Cluster 0 cells 1 cells 2 cells 19 cells 30 cells 72 cells 26 cells

Cluster

Summary statistics of this dataset:
scEx.RData ...
No. of cells: 200
No. of genes: 958
Median UMs per cell: 76.5
Median Genes with min 1 UMI: 65.5
Total number of reads: 14930
Memory used: 493 Mb
Normalization used: DE_logNormalization
Generate report
Download counts.csv
Download Rds

SCHNAPPS

Subclustering: Select a group of cells in plot1 and a different group of cells in plot2 for identifying differential features between these subclusters

colors: colored by cluster identity

selection hint: you can also select by groups you have defined in other plots.

selection hint: also check out "Gene.count" to verify that number genes per cell.

Cluster 1 2 3 4 5 6 X tsne1 Y tsne2

tsne2

tsne1

tsne2

tsne1

Method to use for differential gene expression analysis

Method to use

- Chi-square test of an estimated binomial distribution
- t-test

Differentially Expressed Genes

Selected items to be copied

Download table

Cells

Select all rows

reorder cells by sum of selected genes

By Bernd Jagla (Pasteur Paris)

<https://c3bi-pasteur-fr.github.io/UTechSCB-SCHNAPPS>

<https://github.com/C3BI-pasteur-fr/UTechSCB-SCHNAPPS>

Acknowledgements

Marc Deloger

Morgane Thomas-Chollier

Agnès Paquet

Antonio Rausell

Wouter Saelens

Nathalie Gaspar



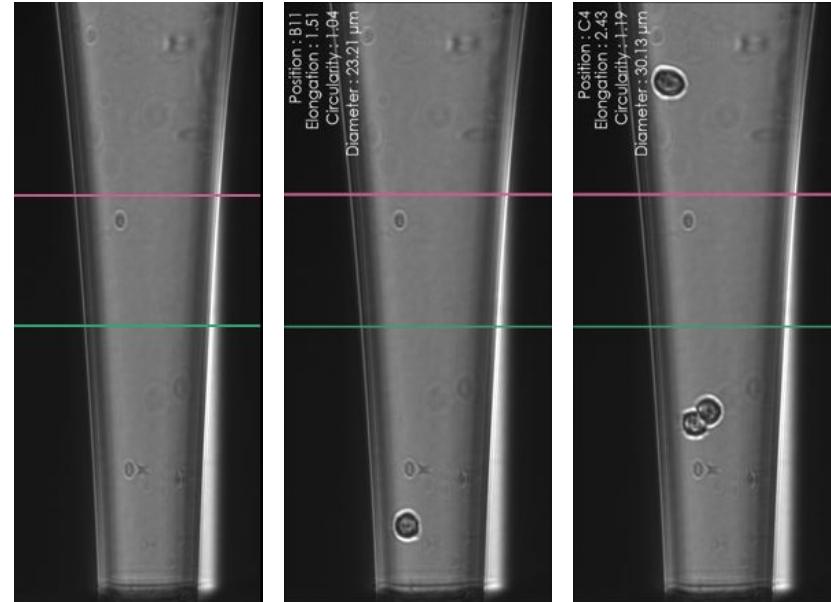
SINGle-cellING in the RAINaseq (1952)*

* play on words courtesy of Jacques van Helden

APPENDIX

Alternative isolation method : Cellenion IBSCI™ (Image Based Single Cell Isolation)

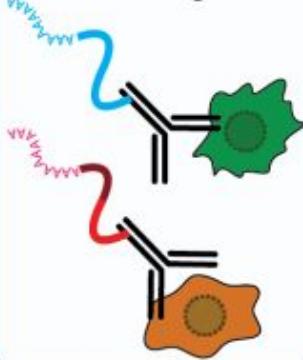
- Capillary real-time **video** recording :
 - Cell or no cell ?
 - More than 1 cell ?
 - Cell size ?
- Acoustic dispersion (more gentle)
- Middle scale :
 - Plate-based
 - Up to 1532 cells
- Cell recovery rate over 95%
- Open platform
 - Scalable, compatible
 - Custom reaction kits



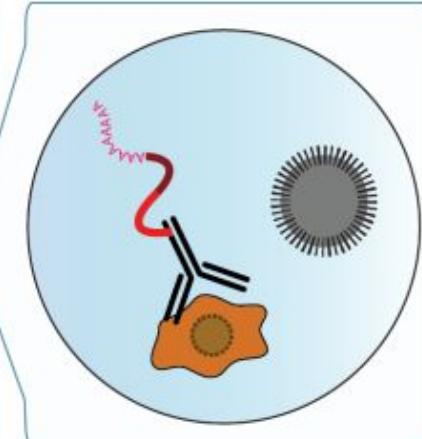
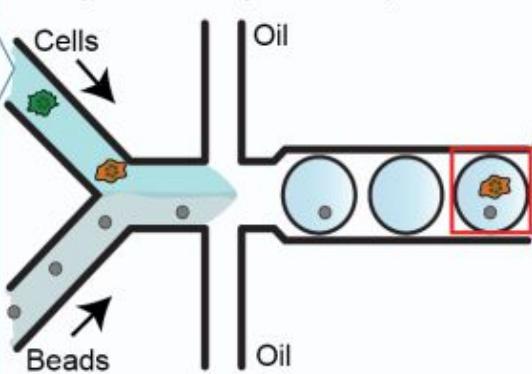
CITE-seq

Cellular Indexing of Transcriptomes and Epitopes by Sequencing

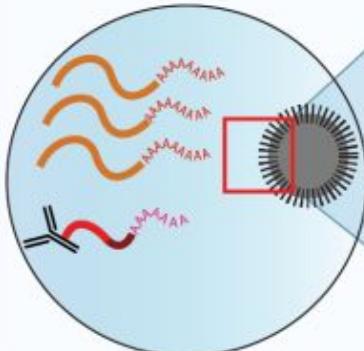
Antibody binding, washing cells



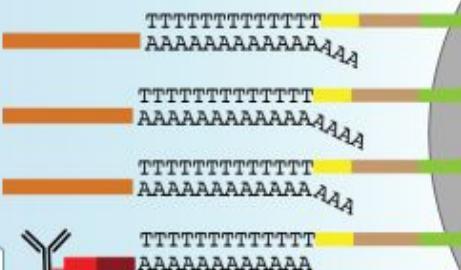
Single cell droplet encapsulation



Cell lysis in droplet



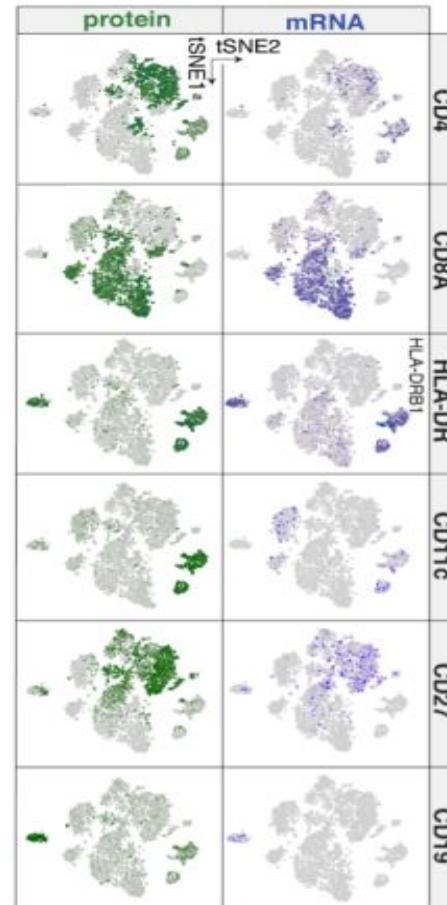
mRNAs and antibody-oligos hybridize to RT oligos and are indexed with cell barcode



Size selected cDNA for standard library prep

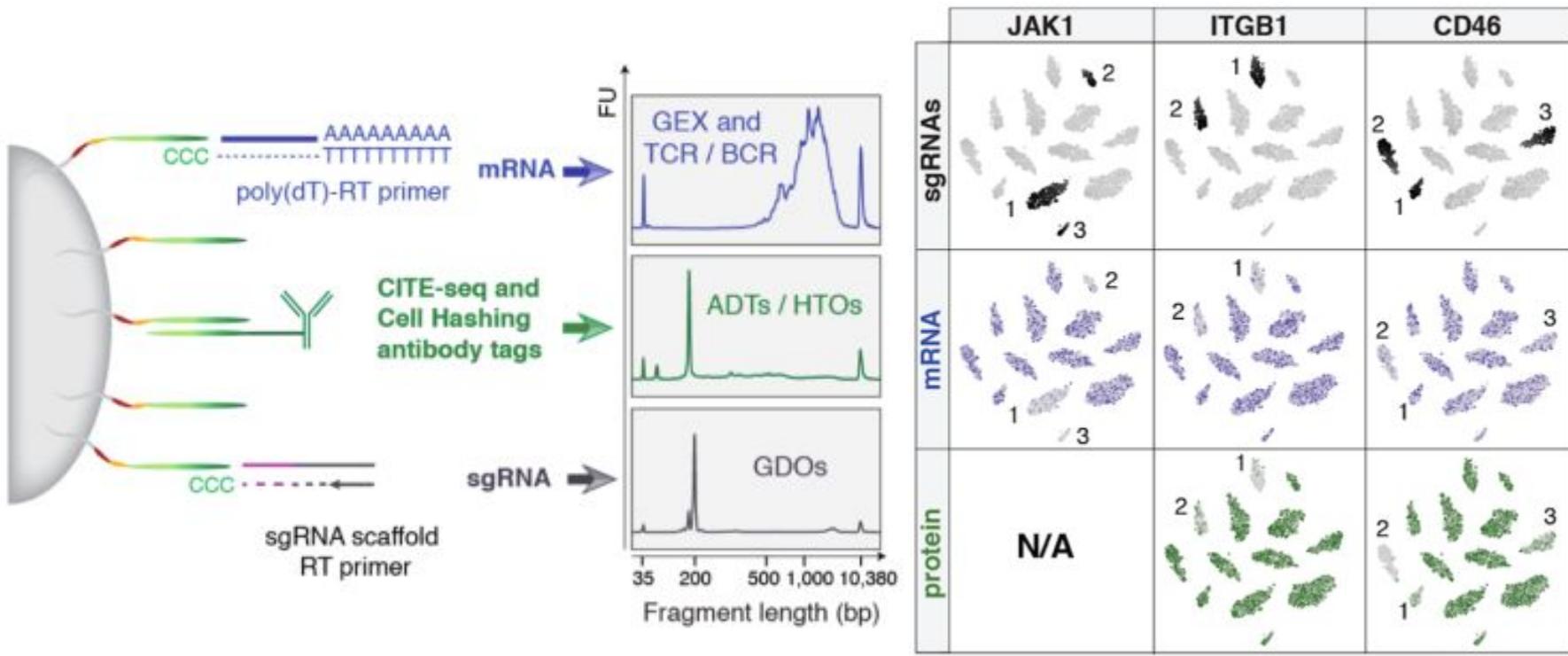


Size selected antibody oligo products for further library prep

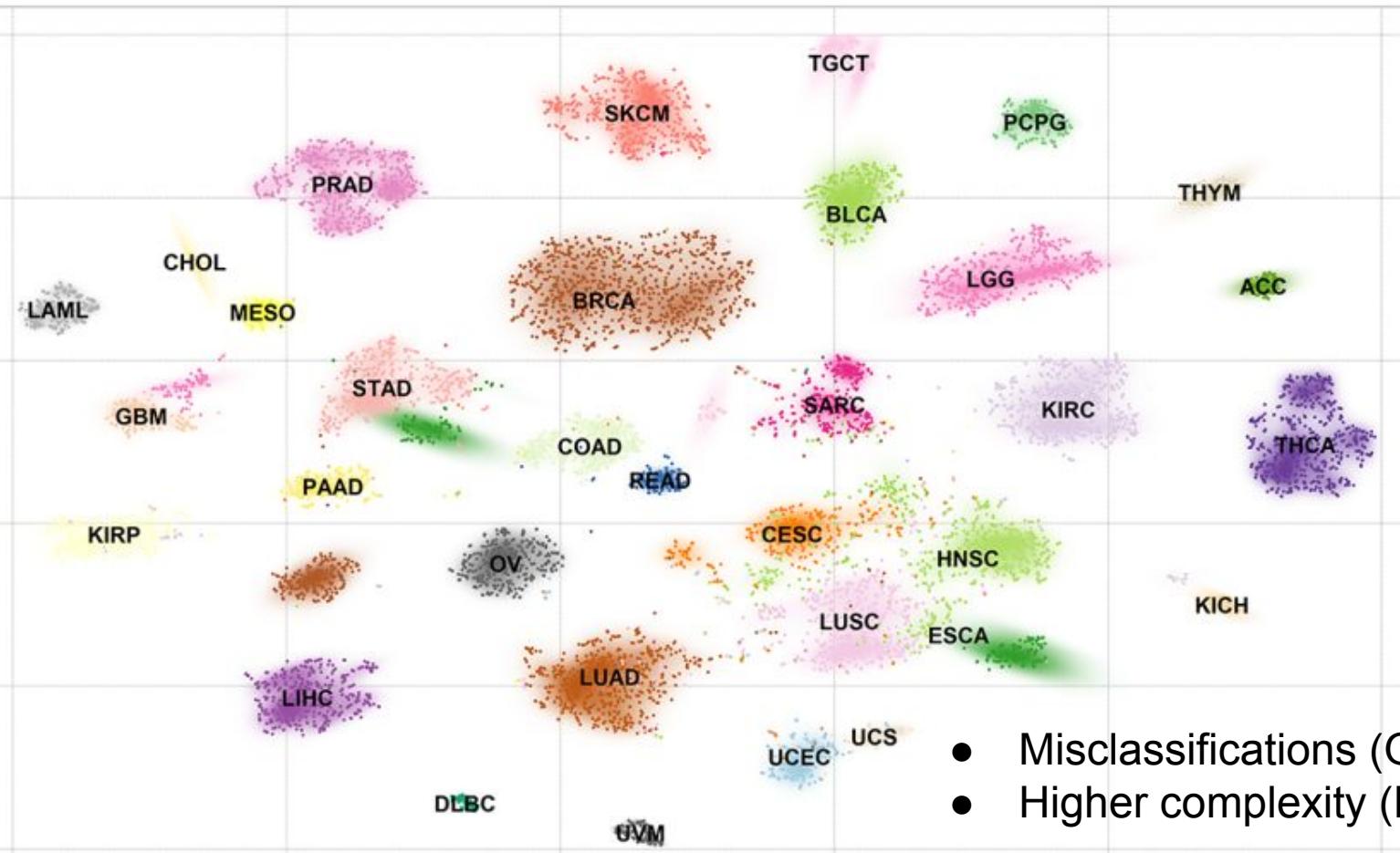


ECCITE-seq

Extended CRISPR-compatible Cellular Indexing of Transcriptomes and Epitopes by Sequencing (5')



t-SNE of the whole TCGA project



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

I. Tissue Procurement

Source:

- Primary human
- Model organism
- Cell culture

Key considerations:

- Biological variation
- Sampling/handling variation
- Duration of sourcing

Study design:

- Biological replicates
- Technical replicates
- Cell number calculation
- Workflow optimization

II. Tissue Dissociation

Method:

- Mechanical mincing
- Enzymatic digestion
- Automated blending
- Microfluidics devices

Key considerations:

- Experimental consistency
- Shortest duration
- Highest cell/nucleus quality
- Representation of all cell types

Quality control:

- FACS analysis
- qPCR for marker genes
- Imaging of cell integrity
- RNA quality (RIN)

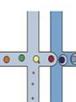
III. Cell Enrichment (optional)

Method:

- Differential centrifugation, sedimentation, filtration
- Antibody labeling for positive/negative selection
- Flow cytometry or bead-based enrichment
- Dead cell removal

Key considerations:

- Additional handling
- Longer duration
- Loss of RNA quality
- Transcriptome changes

IV. Single Cell RNAseq Platform

Method:

- Droplet-based
- Tube-based after FACS
- Microwell-based
- Microfluidics-enabled

Key considerations:

- Cell throughput and handling time
- Gene coverage and cell type detection
- Whole transcript versus 3'end counting
- Imaging capability for doublet detection

V. Library Sequencing

Method:

- Illumina NGS
- Compatible with cDNA library

Sequencing depth considerations:

- 3'end counting: low depth ~50K RPC
- Whole transcript: high depth ~1M RPC
- Alternative splicing: ~20-30M RPC
- Iterative optimization for biological system

VI. Computational Analysis

Key considerations:

- Separation of batch and condition
- Technical vs. biological variation

Sample Batch correction approaches:

- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST