

Introduction to single-cell ‘omics

Tomás Gomes, PhD

29 April 2024



@tomsgoms

tomas.gomes@medicina.ulisboa.pt

Applied Computational Multi-Omics 23/24

FCT UNL

Introduction

FCUL, Lisboa



2009 - BSc Biology
MSc Comp Biology and Bioinformatics

WSI, Cambridge



2016 - PhD (use scRNA-seq to study cross tissue cell type differences)

iMM, Lisboa



2023 – Post-doc (immunology with scRNA-seq and spatial transcriptomics)

iMM, Lisboa



2014 - MSc thesis (Mapping nascent transcripts to the genome)

ETHZ, Basel



2020 - Post-doc (using single-cell technologies to study regeneration)

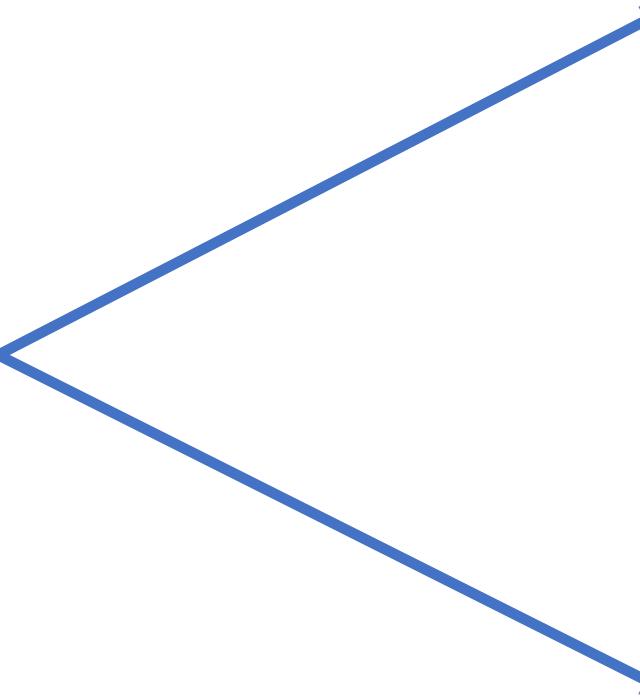
Topics:

- The emergence of single-cell sequencing
- How to get mRNA from single-cells
- Processing and analysis steps
- Data integration and multi-modal data
- Areas of intense, active development

Topics:

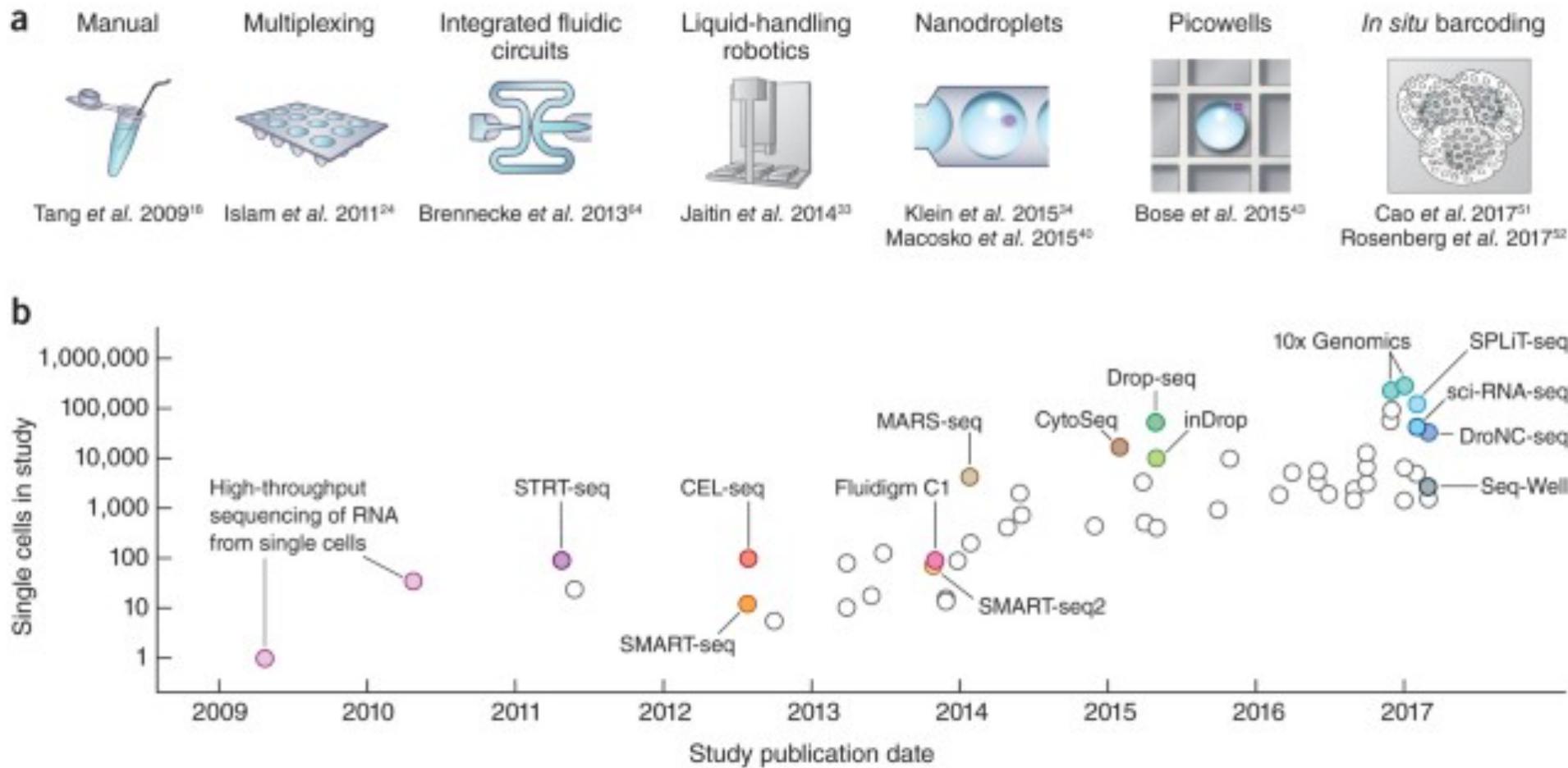
- The emergence of single-cell sequencing
- How to get mRNA from single-cells
- Processing and analysis steps
- Data integration and multi-modal data
- Areas of intense, active development

Why single-cell?



JUSTATASTE.COM

The single-cell sequencing decade



The first instance of single-cell resolution RNA-seq

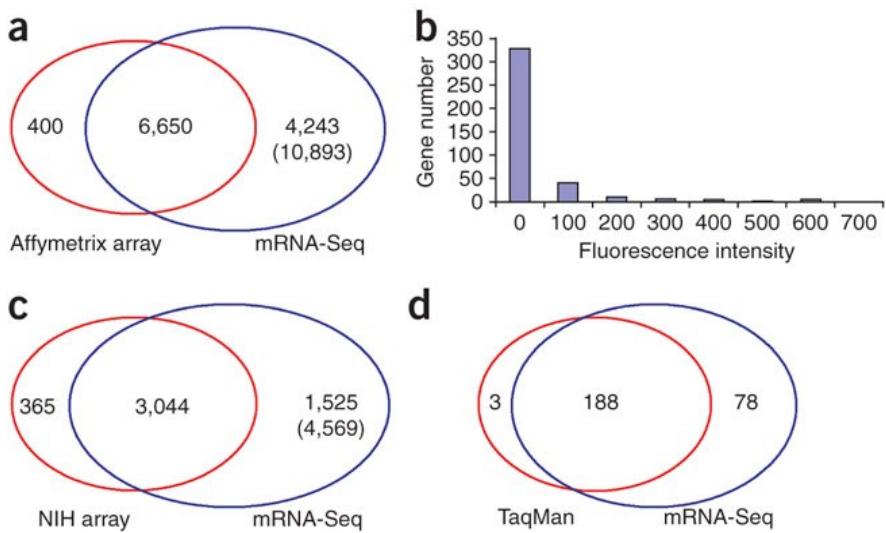
Published: 06 April 2009

mRNA-Seq whole-transcriptome analysis of a single cell

Fuchou Tang, Catalin Barbacioru, Yangzhou Wang, Ellen Nordman, Clarence Lee, Nanlan Xu, Xiaohui Wang, John Bodeau, Brian B Tuch, Asim Siddiqui, Kaiqin Lao✉ & M Azim Surani✉

Nature Methods **6**, 377–382 (2009) | [Cite this article](#)

14k Accesses | **1308** Citations | **89** Altmetric | [Metrics](#)



1 year after the [first RNA-seq paper](#)

Goal was to quantify RNA in a very low input system (but with many different genes expressed)

Setting the stage for the single-cell decade

Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq

Saiful Islam^{1,4}, Una Kjällquist^{1,4}, Annalena Moliner², Paweł Zajac¹, Jian-Bing Fan³, Peter Lönnerberg¹ and Sten Linnarsson^{1,5}

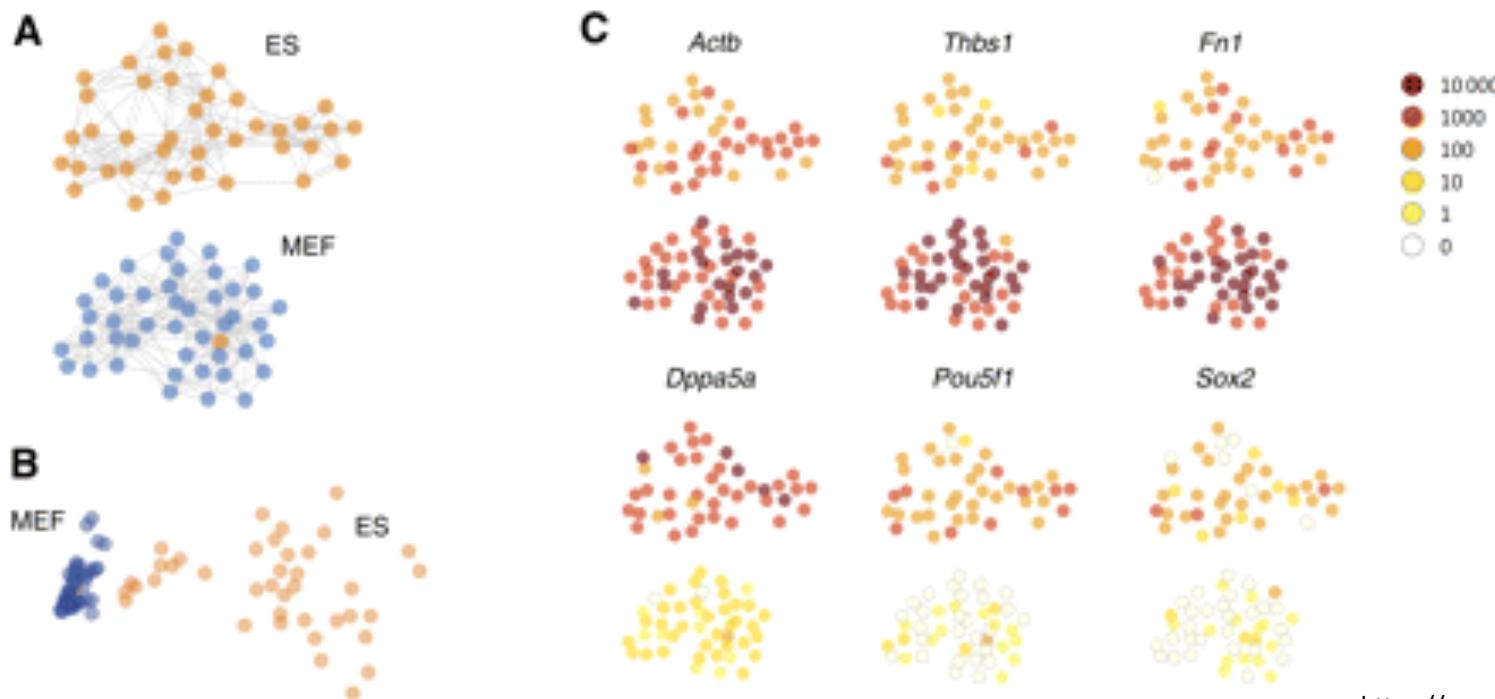
¹Laboratory for Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-171 77 Stockholm, Sweden;

²Department of Neuroscience, Karolinska Institutet, SE-171 77 Stockholm, Sweden;

³Illumina Inc., San Diego, California 92121, USA

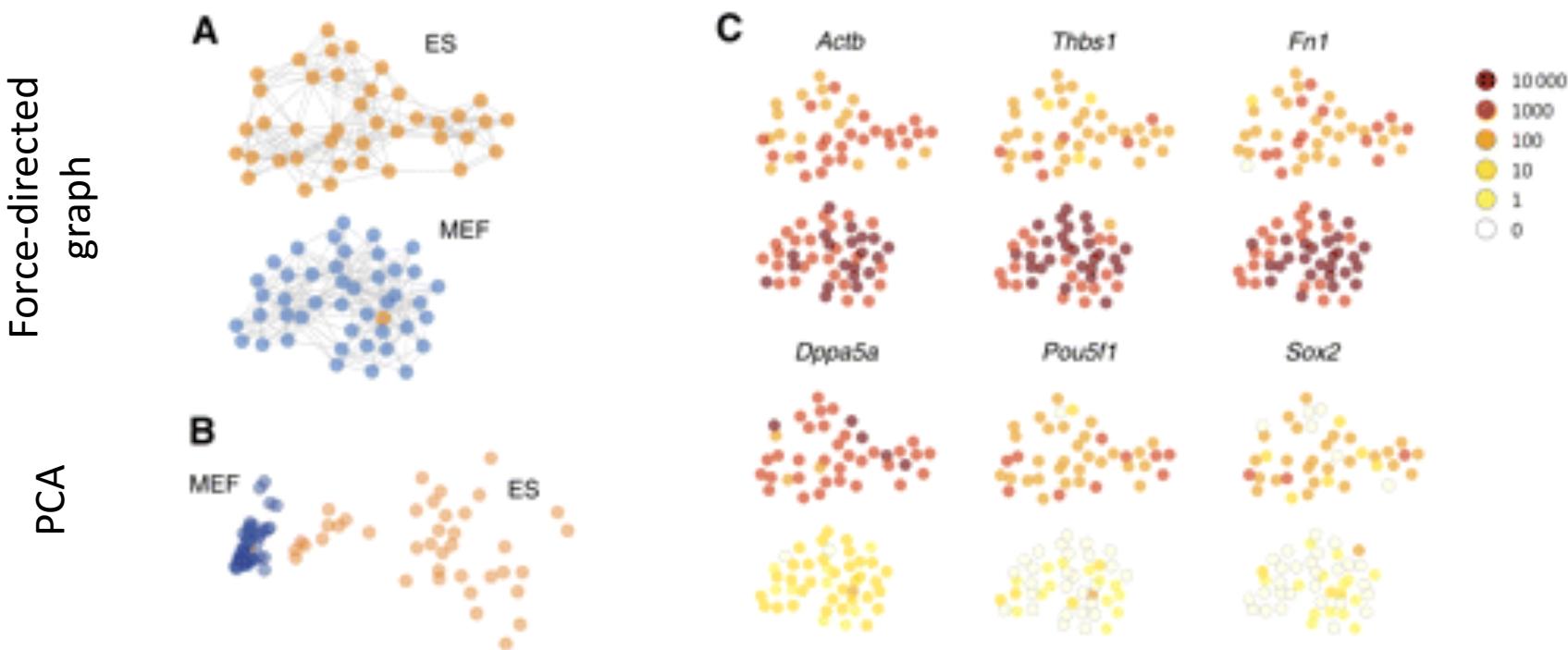
First paper (2011) doing multiplexing of single-cells in a 96 well plate

Use of **control RNA molecules** (ERCC, included at specific concentrations) to distinguish technical and biological variance



Setting the stage for the single-cell decade

We envisage the future use of very large-scale single-cell transcriptional profiling to build a detailed map of naturally occurring cell types, which would give unprecedented access to the genetic machinery active in each type of cell at each stage of development.



Topics:

- The emergence of single-cell sequencing
- How to get mRNA from single-cells
- Processing and analysis steps
- Data integration and multi-modal data
- Areas of intense, active development

Common steps in a scRNA-seq experiment

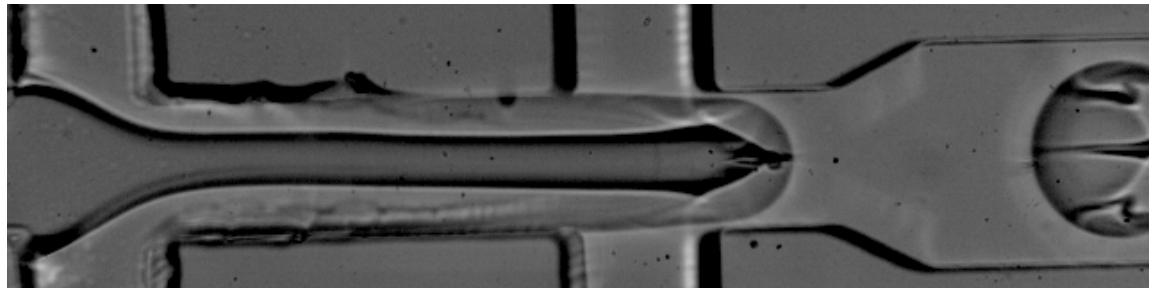
scRNA-seq experiment	<u>Bench work</u>	Cell isolation
		Library construction
		Sequencing
		Quantification
	<u>Computational work</u>	Quality control
		Normalisation
		Dimensionality reduction
		Clustering/Trajectory analysis
		Differential expression

Isolation - how do we get single-cells?

Plate-based
methods



Microdroplet-based
methods



Goal is to be able to perform RNA extraction and library construction for each cell in isolation

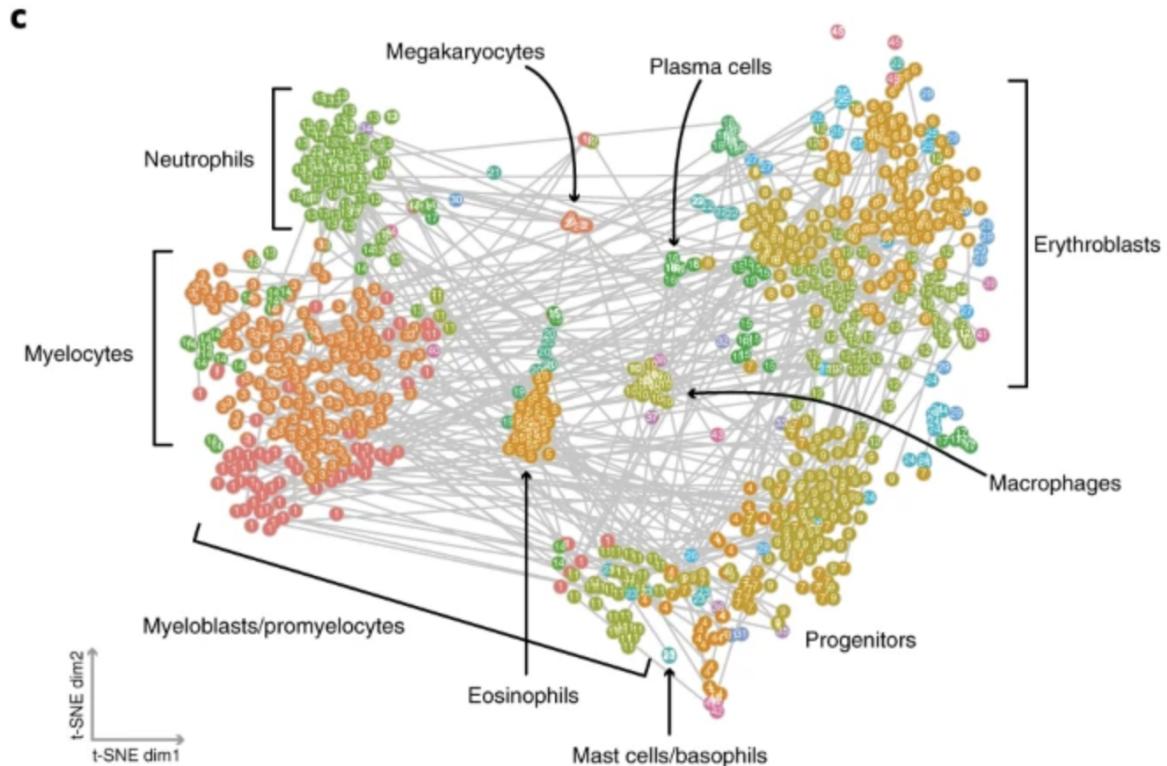
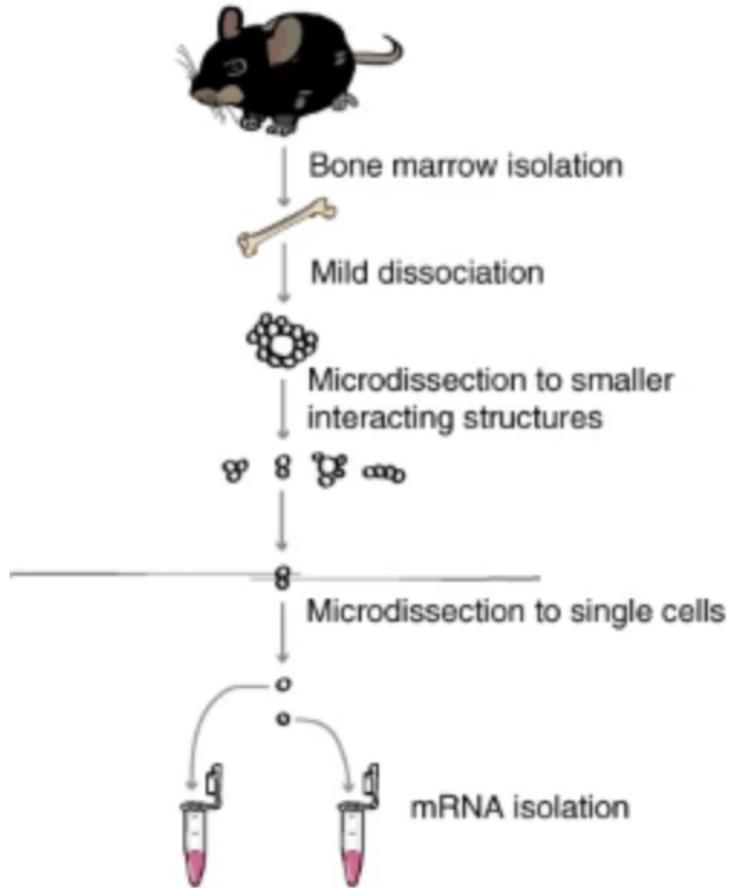
Microdissection: uncommon but useful

Article | Published: 21 May 2018

Mapping the physical network of cellular interactions

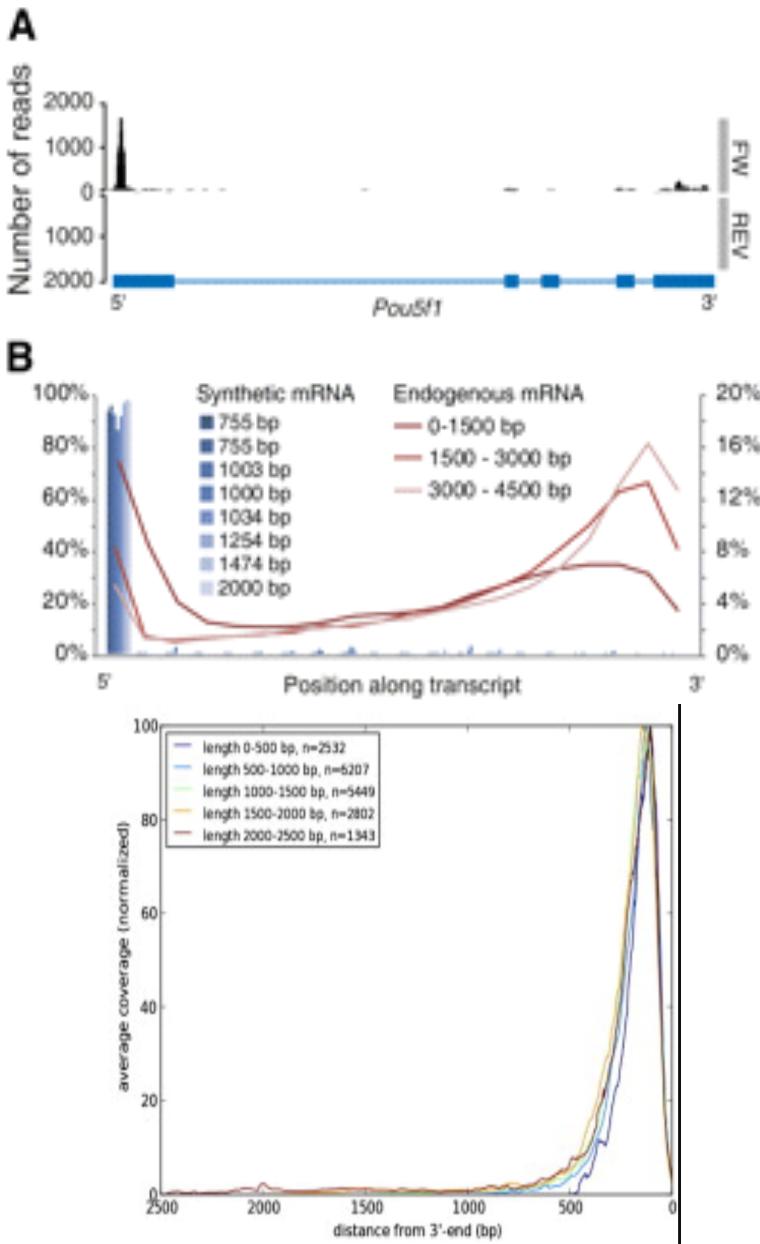
Jean-Charles Boisset, Judith Vivié, Dominic Grün, Mauro J. Muraro, Anna Lyubimova & Alexander van Oudenaarden [✉](#)

Nature Methods 15, 547–553 (2018) | [Cite this article](#)



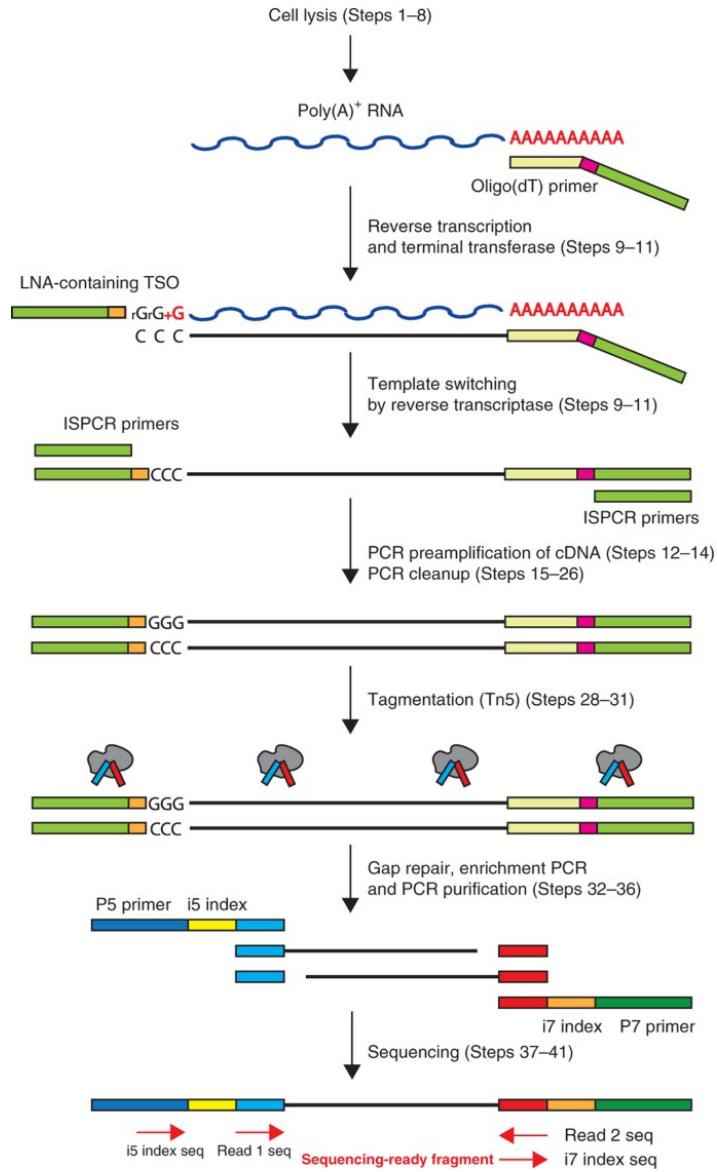
Microdissection of interacting cell pairs to **identify preferentially interacting cell types**

Isolating and amplifying RNA from single-cells



- 5' end capture (e.g. STRT-seq)
- 3' end capture (e.g. CEL-seq)
- Whole transcript capture (e.g. Smart-seq2)

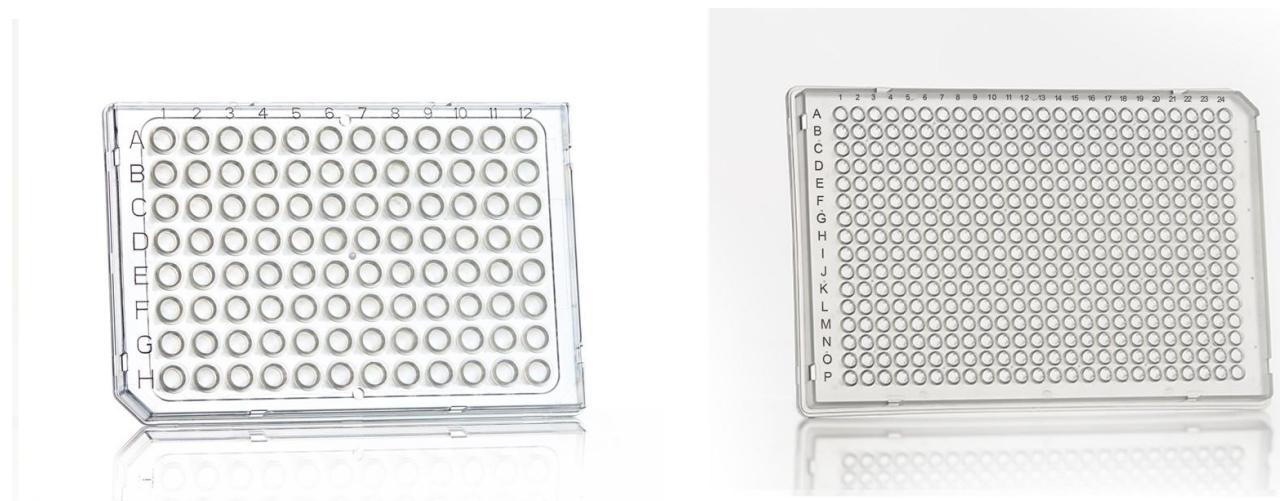
Getting RNA from single-cells - Smart-seq2



Smart-seq (current version V4, the most used/well known is V2) is a protocol used to sequence the **full-length** transcriptome of individual cells

Requires extensive liquid handling, usual **throughput is ~1000 cells**

Since fewer cells are sequenced, sequencing depth tends to be higher and thus more genes are captured



Getting RNA from single-cells - Chromium

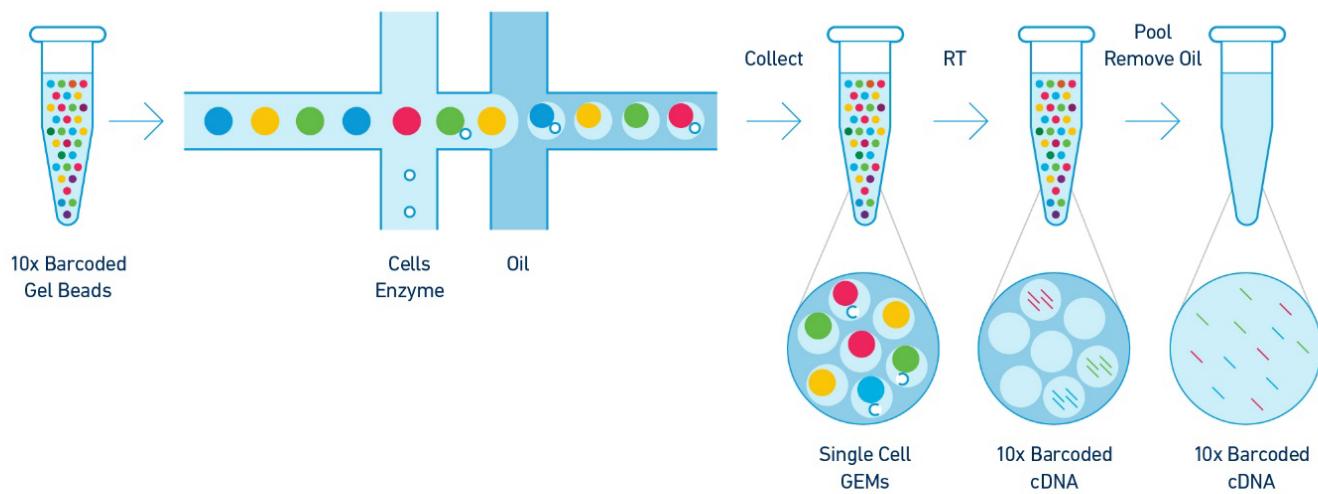
Open Access | Published: 16 January 2017

Massively parallel digital transcriptional profiling of single cells

Grace X. Y. Zheng, Jessica M. Terry, Phillip Belgrader, Paul Ryvkin, Zachary W. Bent, Ryan Wilson, Solongo B. Ziraldo, Tobias D. Wheeler, Geoff P. McDermott, Junjie Zhu, Mark T. Gregory, Joe Shuga, Luz Montesclaros, Jason G. Underwood, Donald A. Masquelier, Stefanie Y. Nishimura, Michael Schnall-Levin, Paul W. Wyatt, Christopher M. Hindson, Rajiv Bharadwaj, Alexander Wong, Kevin D. Ness, Lan W. Beppu, H. Joachim Deeg, Christopher McFarland, Keith R. Loeb, William J. Valente, Nolan G. Ericson, Emily A. Stevens, Jerald P. Radich, Tarjei S. Mikkelsen, Benjamin J. Hindson  & Jason H. Bielas  - Show fewer authors

Nature Communications 8, Article number: 14049 (2017) | Cite this article

67k Accesses | 1210 Citations | 201 Altmetric | Metrics



10x Genomics Chromium platform

Uses microdroplet encapsulation for individual lysis and amplification

Current version is v3.1; has 3' and 5' sequencing variants

High throughput (5-10k cells/well)

Sequencing more than just mRNA

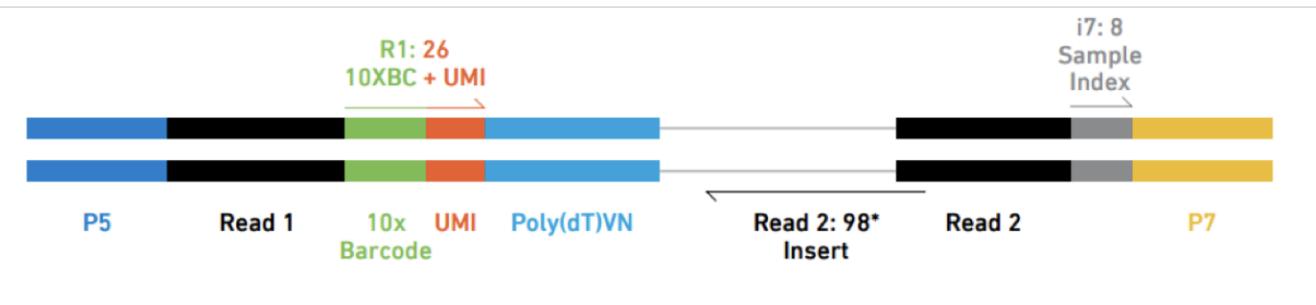


Fig. 2. Schematic of a fragment from a final Chromium™ Single Cell 3' v2 library. *Can be adjusted.

For highly multiplexed experiments, transcripts from each cell should be tagged with a **cell barcode** to allow sequencing in bulk and further deconvolution

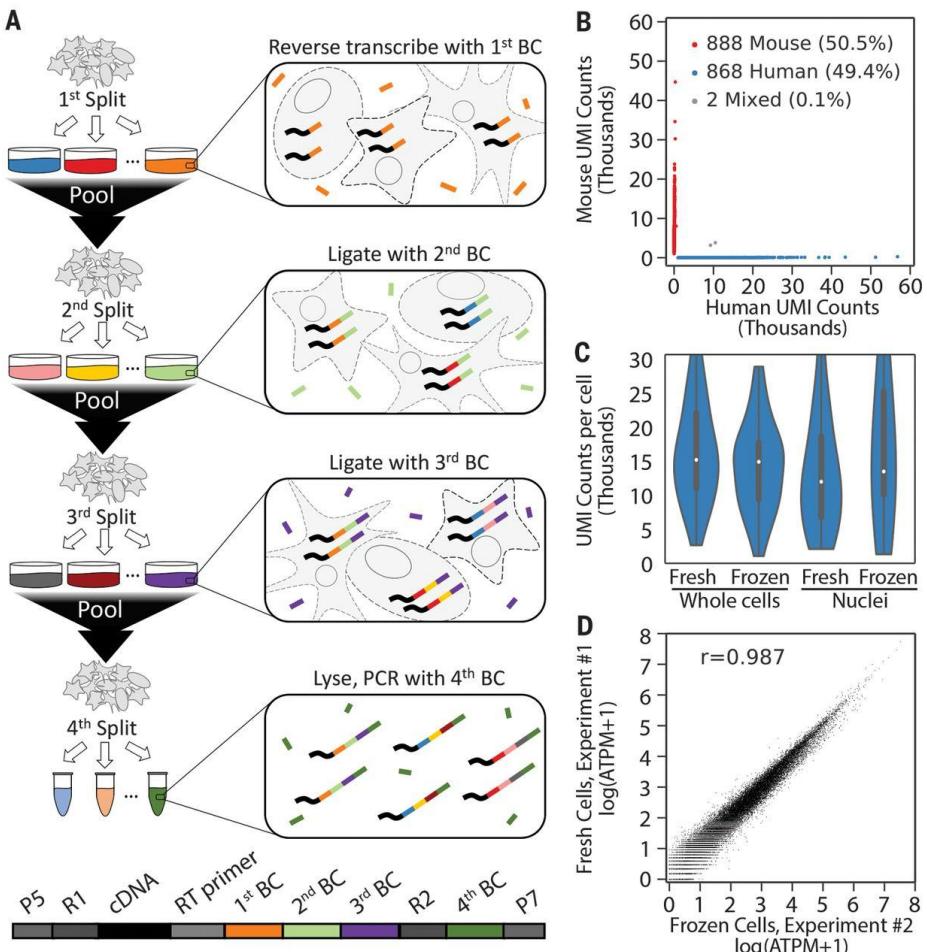
To avoid counting the same transcript twice (usually caused by PCR duplicates), a Unique Molecular Identifier (UMI) is also added to the constructed libraries

This also means that counting UMI duplication can be a QC metric for the sample/sequencing

The return of plate-based methods?

Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding

✉ Alexander B. Rosenberg^{1,*†}, □ Charles M. Roco^{2,*}, □ Richard A. Muscat¹, □ Anna Kuchina¹, □ Paul Sample¹, □ Zizhen Yao³, □ Lucas T. Graybuck³, □ David J. Peeler², □ Sumit Mukherjee¹, □ Wei Chen⁴, □ Suzie H. Pun², □ Drew L. Sellers^{2,5}, □ Bosiljka Tasic³, □ Georg Seelig^{1,4,6,†}



Pooling and splitting cells, and adding consecutive barcodes, will give a very close to unique cell barcoding

This allows for greater throughput with **lower costs**, as well as **no specialised instrument** required (just a 24/48/96 well plate)

Methods: SPLiT-seq, sci-RNA-seq

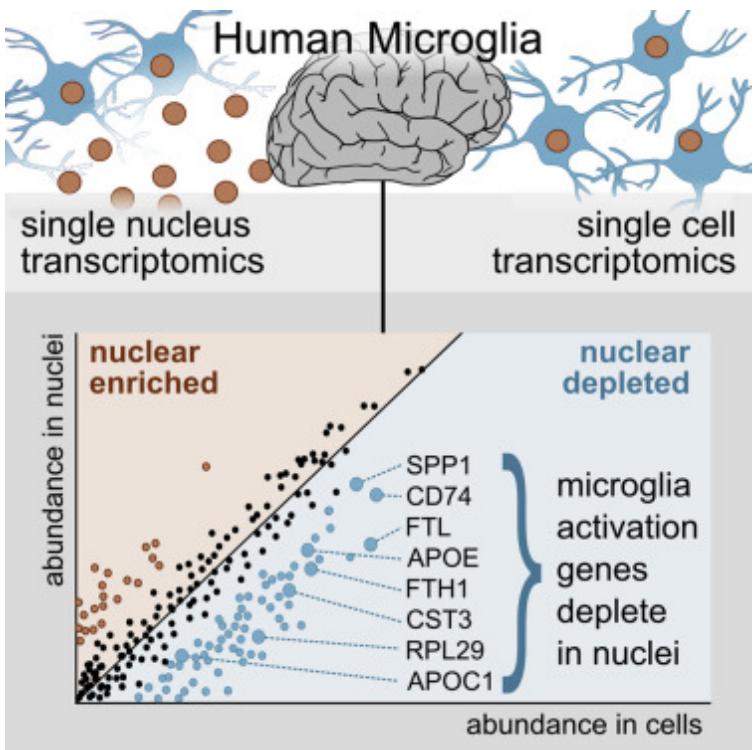


Single-cell vs single-nuclei

Report

Single-Nucleus RNA-Seq Is Not Suitable for Detection of Microglial Activation Genes in Humans

Nicola Thrupp ^{1, 2}, Carlo Sala Frigerio ^{1, 2, 3}, Leen Wolfs ^{1, 2}, Nathan G. Skene ⁴, Nicola Fattorelli ^{1, 2}, Suresh Poovathingal ^{1, 2}, Yannick Fourne ^{1, 2}, Paul M. Matthews ⁴, Tom Theys ⁵, Renzo Mancuso ^{1, 2}, Bart de Strooper ^{1, 2, 3, 6}                                                  



In some situations, it might be necessary to sequence **single-nuclei** instead of single-cells

Applications: frozen samples, oddly-shaped cells (e.g. neurons)

Usually RNA content and gene diversity is lower than whole cell, but still possible to identify cell populations

Transcripts in the nucleus are not *exactly* the same as those in the cytoplasm

Topics:

- The emergence of single-cell sequencing
- How to get mRNA from single-cells
- Processing and analysis steps
- Data integration and multi-modal data
- Areas of intense, active development

Common steps in a scRNA-seq experiment

scRNA-seq experiment	<u>Bench work</u>	Cell isolation
		Library construction
		Sequencing
		Quantification
	<u>Computational work</u>	Quality control
		Normalisation
		Dimensionality reduction
		Clustering/Trajectory analysis
		Differential expression

Quantification and demultiplexing - CellRanger

Open Access | Published: 16 January 2017

Massively parallel digital transcriptional profiling of single cells

Grace X. Y. Zheng, Jessica M. Terry, Phillip Belgrader, Paul Ryvkin, Zachary W. Bent, Ryan Wilson, Solongo B. Ziraldo, Tobias D. Wheeler, Geoff P. McDermott, Junjie Zhu, Mark T. Gregory, Joe Shuga, Luz Montesclaros, Jason G. Underwood, Donald A. Masquelier, Stefanie Y. Nishimura, Michael Schnall-Levin, Paul W. Wyatt, Christopher M. Hindson, Rajiv Bharadwaj, Alexander Wong, Kevin D. Ness, Lan W. Beppu, H. Joachim Deeg, Christopher McFarland, Keith R. Loeb, William J. Valente, Nolan G. Ericson, Emily A. Stevens, Jerald P. Radich, Tarjei S. Mikkelsen, Benjamin J. Hindson & Jason H. Bielas - Show fewer authors

Nature Communications 8, Article number: 14049 (2017) | [Cite this article](#)

67k Accesses | 1210 Citations | 201 Altmetric | [Metrics](#)

A command-line software suite to quantify (and analyse) data from 10x Genomics scRNA-seq/snRNA-seq

Based on **STAR** for alignment

Parses Cell barcodes and UMIs

Detects real cells and empty droplets

Performs some default analysis



Products Research Areas Resources Support Company

Support > Single Cell Gene Expression > Software

SEARCH

Q&A NEW

CONTACT SUPPORT

SOFTWARE > PIPELINES

CELL RANGER

Introduction

- [What is Cell Ranger?](#)
- [What is Feature Barcode Data?](#)

What is Cell Ranger?

Cell Ranger is a set of analysis pipelines that process Chromium single-cell data to align reads, generate feature-barcode matrices, perform clustering and other secondary analysis, and more. Cell Ranger includes four pipelines relevant to the 3' Single Cell Gene Expression Solution and related products:

```
$ cellranger count --id=sample345 \
--transcriptome=/opt/refdata-gex-GRCh38-2020-A \
--fastqs=/home/jdoe/runs/HAWT7ADXX/outs/fastq_path \
--sample=mysample \
--expect-cells=1000 \
--localcores=8 \
--localmem=64
```

<https://www.nature.com/articles/ncomms14049>

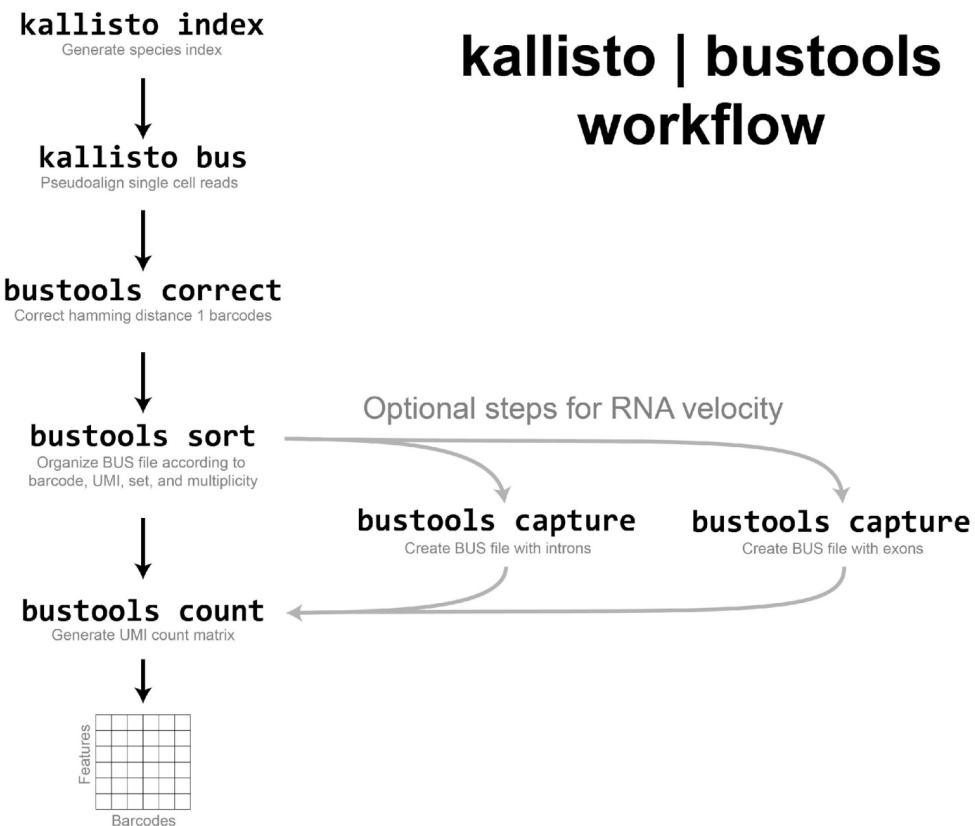
<https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger>

Quantification and demultiplexing – kallisto-bustools

Letter | Published: 01 April 2021

Modular, efficient and constant-memory single-cell RNA-seq preprocessing

Páll Melsted, A. Sina Booeshaghi, Lauren Liu, Fan Gao, Lambda Lu, Kyung Hoi (Joseph) Min, Eduardo da Veiga Beltrame, Kristján Eldjárn Hjörleifsson, Jase Gehring & Lior Pachter 



Kallisto does pseudoalignment of reads to the transcriptome

This is more efficient – faster, low memory and CPU footprint (should work on a standard laptop!)

Kallisto individually is sufficient for plate-based scRNA-seq

BUSTools processes/deconvolutes scRNA-seq data from microdroplet methods

Salmon-Alevin are similar methods

Sequencing more than just mRNA

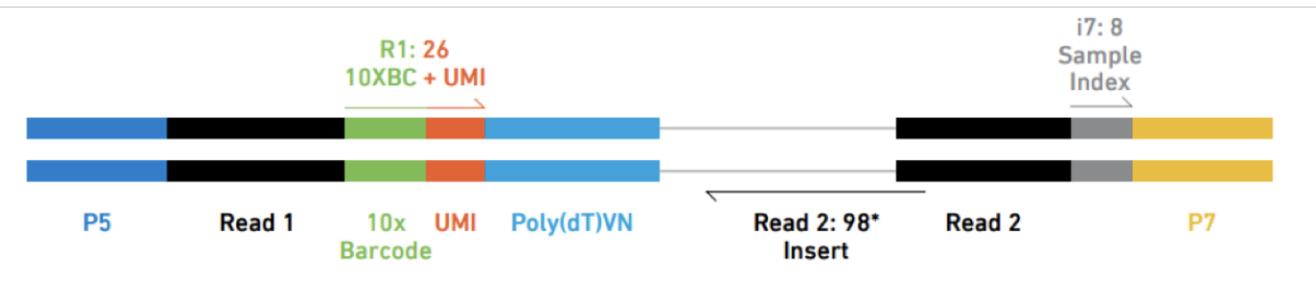


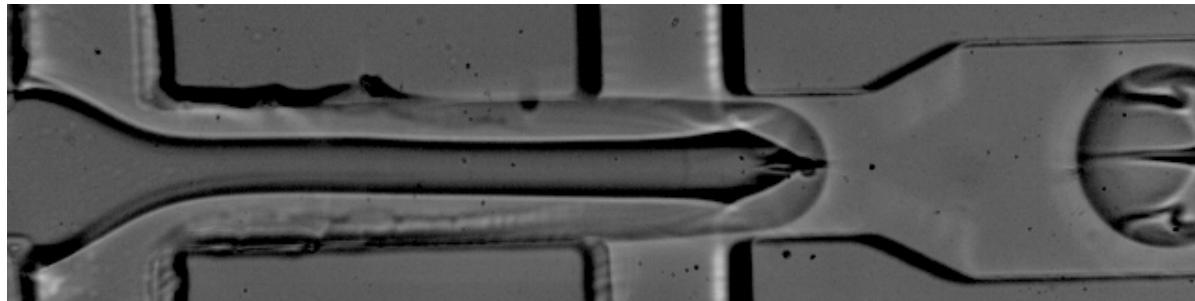
Fig. 2. Schematic of a fragment from a final Chromium™ Single Cell 3' v2 library. *Can be adjusted.

For highly multiplexed experiments, transcripts from each cell should be tagged with a **cell barcode** to allow sequencing in bulk and further deconvolution

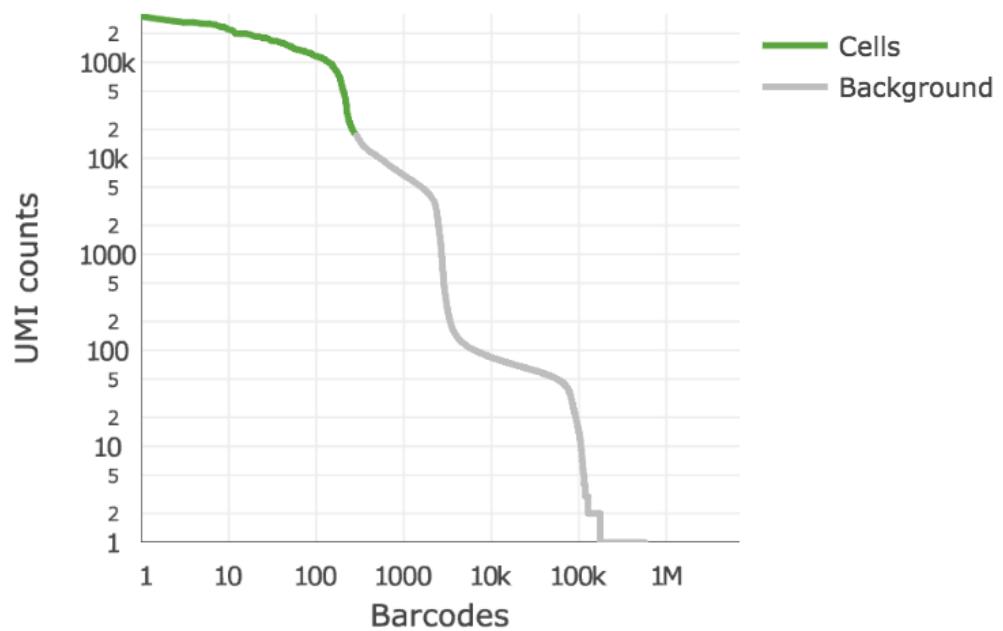
To avoid counting the same transcript twice (usually caused by PCR duplicates), a Unique Molecular Identifier (UMI) is also added to the constructed libraries

This also means that counting UMI duplication can be a QC metric for the sample/sequencing

What is in a microdroplet?



*this is from Drop-seq, but in the Chromium the principle is similar



Droplets can:

- contain nothing – **empty droplet**
- contain RNA but no cell – **empty with ambient RNA**
- contain a cell - **cell**
- contain more than one cell - **doublet**

Specific methods exist to distinguish these, mostly based on:

- UMI counts
- Transcriptomic similarity with empty droplets
- Transcriptomic similarity between cell populations

UMI counts and number of genes are thus important QC metrics

- High counts can indicate doublets
- Low counts can indicate dead cells/empty droplets

Specific gene sets are also important in QC

- **MT genes** – high in dying cells
- **Stress genes** (Heat Shock Proteins, ...) – tissue dissociation effects

Importance of normalisation

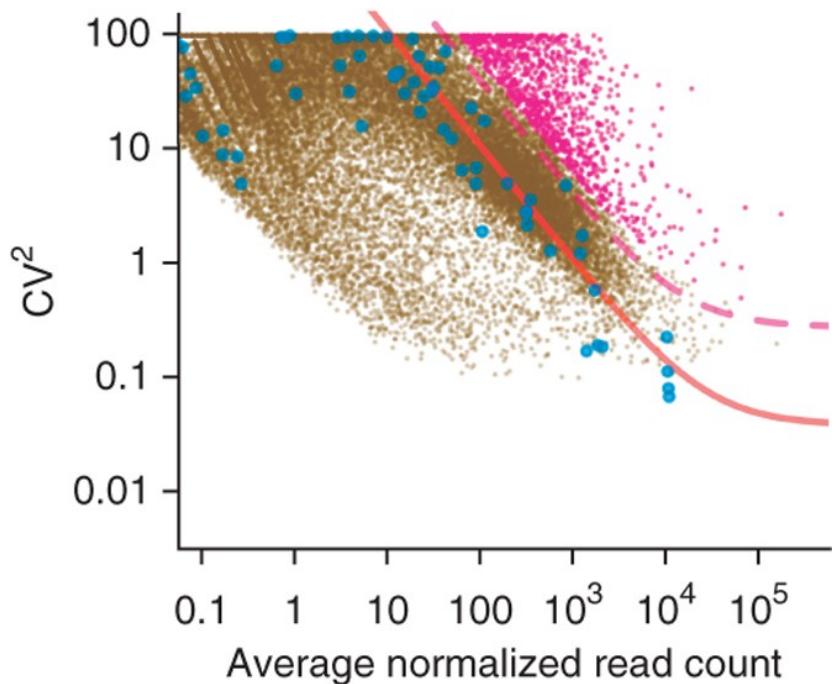
- The most common way to normalise scRNA-seq data is **dividing UMI counts by the total per cell**, followed by a **log-transformation**
- A “pseudocount” has to be added to non-expressed genes (since you can’t have $\log(0)$) – usually 1, but ideally a fraction (0.1 or lower)
- In this step, **regressing out the effect of other technical/undesired variables** (e.g. total counts per cell, % MT genes, cell cycle) can also be performed
- Further **scaling** (i.e. subtract the mean (and optionally divide by the standard deviation)) ensures the data can be used for linear projections (e.g PCA)

Obtaining “meaningful” genes

Published: 22 September 2013

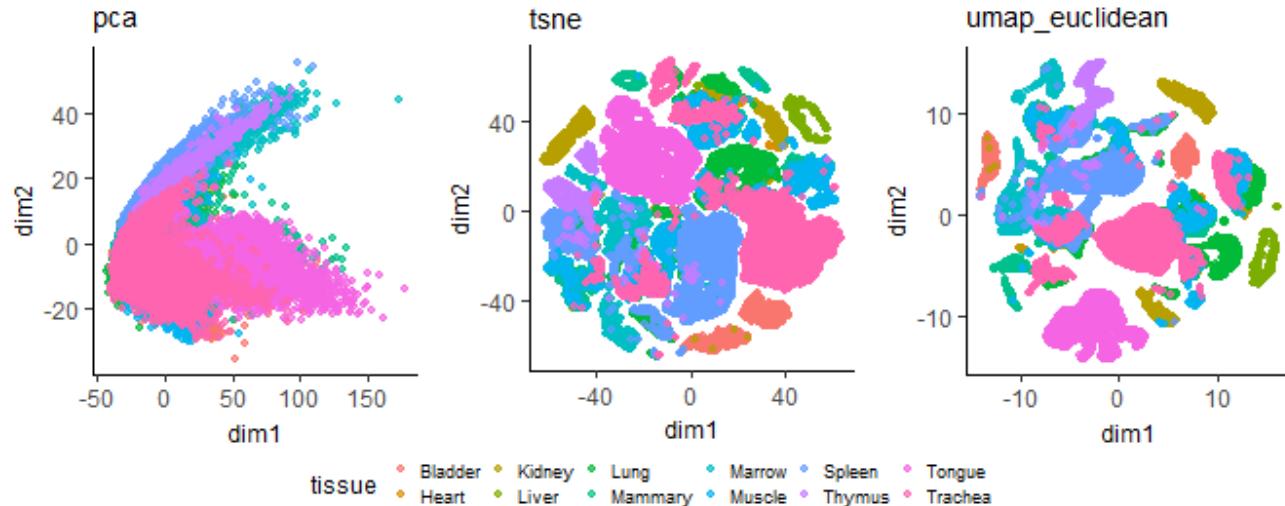
Accounting for technical noise in single-cell RNA-seq experiments

Philip Brennecke, Simon Anders, Jong Kyoung Kim, Aleksandra A Kołodziejczyk, Xiuwei Zhang, Valentina Proserpio, Bianka Baying, Vladimir Benes, Sarah A Teichmann, John C Marioni & Marcus G Heisler

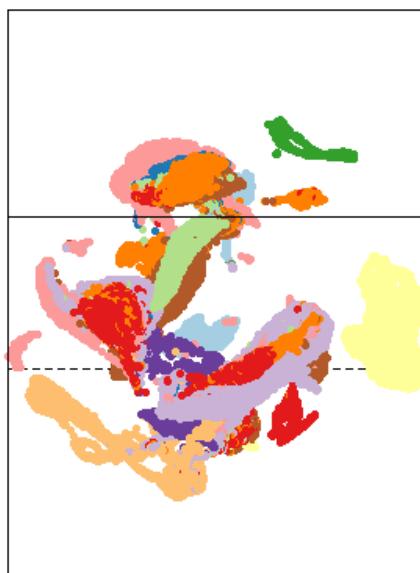


- Genes don't all equally reflect biological variance. For example, housekeeping genes should be ubiquitously expressed, and therefore not be useful indistinguish cell populations
- Additionally, reducing the feature (gene) space in scRNA-seq data helps speed up computation
- This gene subset should *only* be used for dimensionality reduction/clustering/trajectory analysis, not the characterization of cell populations (i.e. differential expression)
- Ultimately, there are various ways to obtain them, and those should be tuned to the dataset at hand

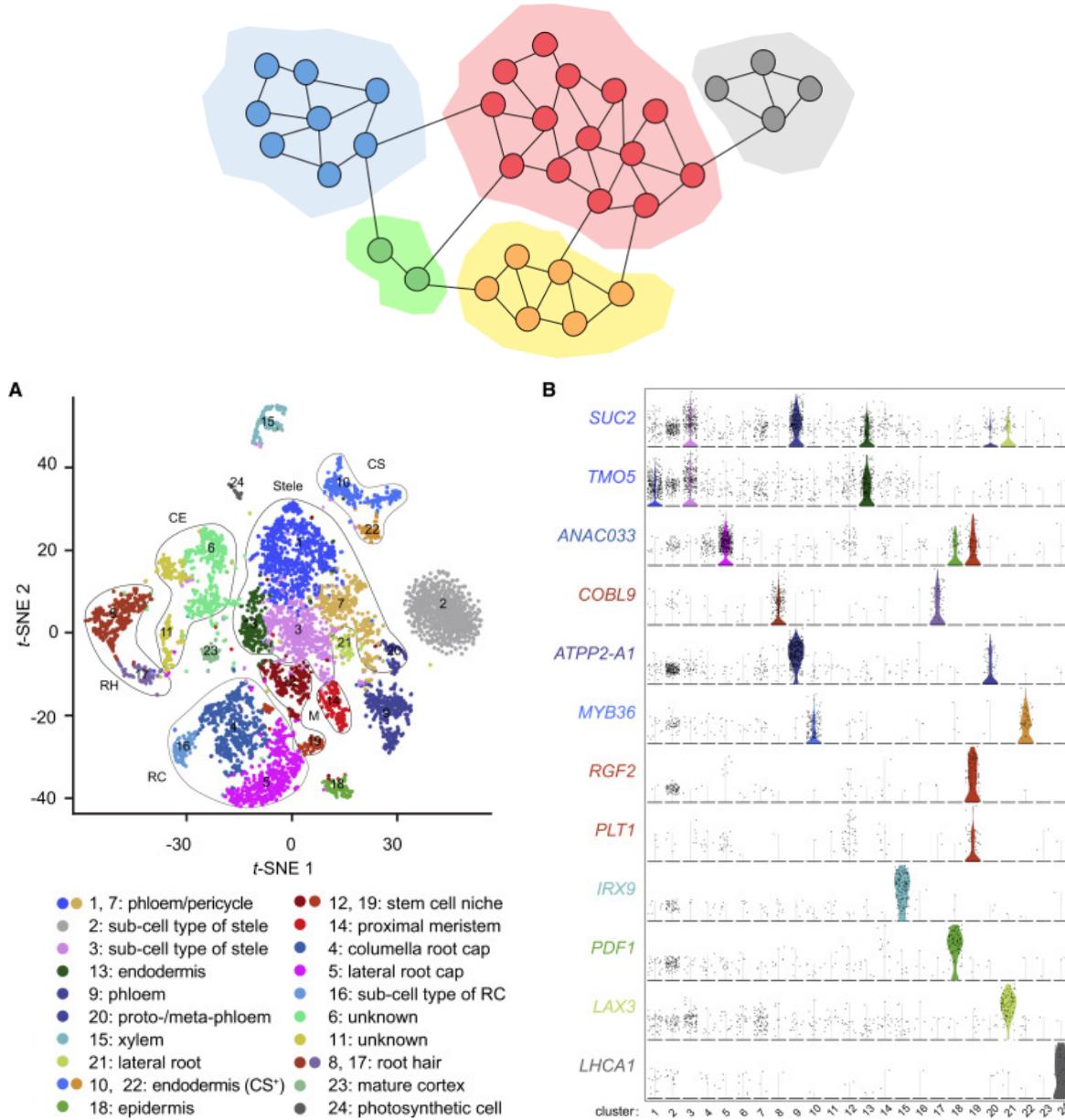
Dimensionality reduction



- scRNA-seq datasets commonly have ~ 10.000 *data points* (cells) and ~ 20.000 *features* (genes)
 - hard to visualise!
- DR algorithms help us visualize the major variability within a dataset
- PCA is the most common linear DR method, whereas tSNE and UMAP are non-linear
- PCA gives us orthogonal dimensions reflecting different aspects of the dataset's variability
- tSNE and UMAP are used to “flatten” the dataset into a 2D representation, **not to do any further analysis**, since they distort distances between cells



Identifying cell types from scRNA-seq - clustering



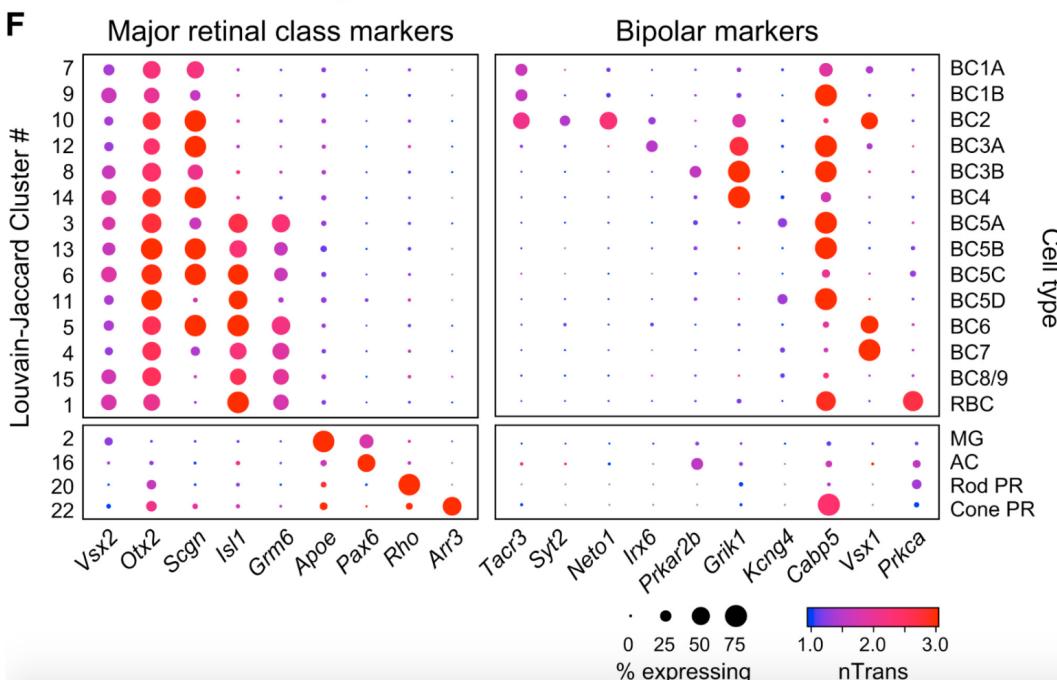
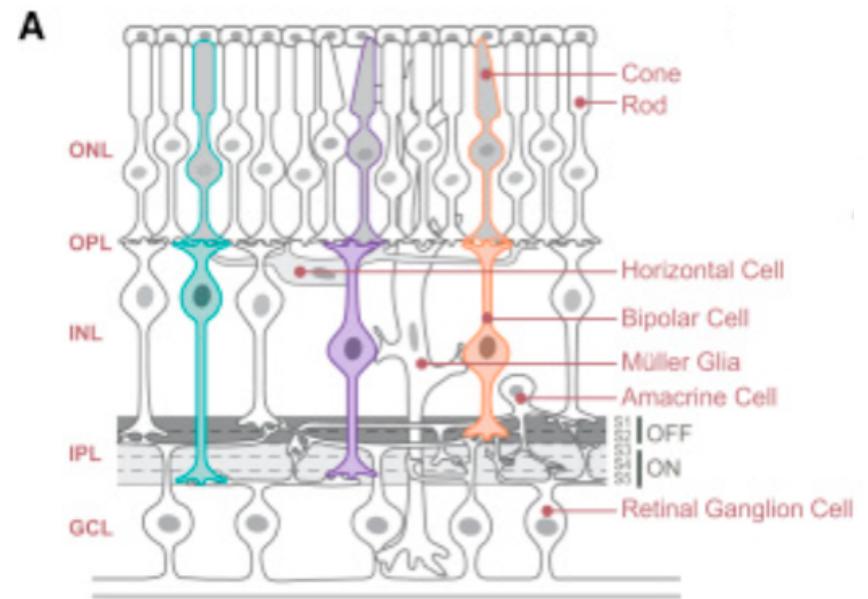
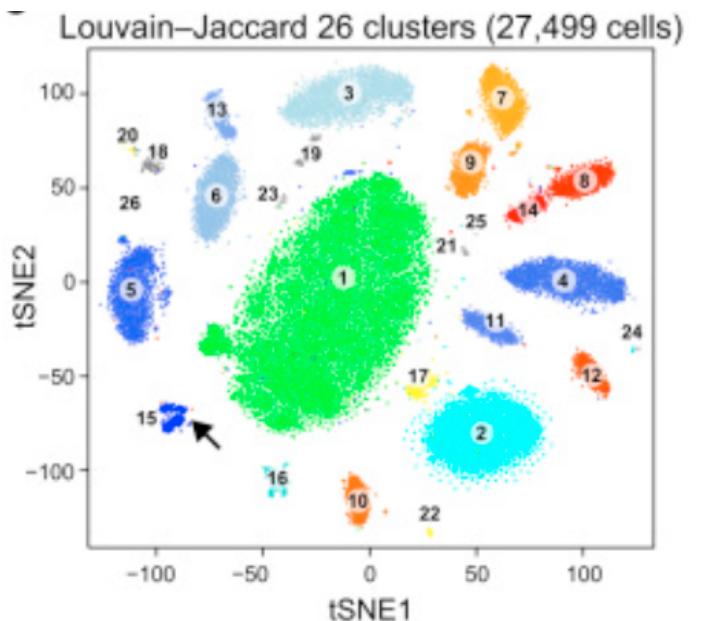
- Clustering is used to **identify cell populations**
- These can be interpreted based on their marker genes to define cell types
- Marker genes are obtained with differential expression testing of one population vs the rest
- The most popular clustering methods, **Leiden** and **Louvain**, are **community detection methods**
- They work on **cell neighbourhood graphs**, usually computed from **PCA space, NOT UMAP/tSNE**

Identifying cell types from scRNA-seq

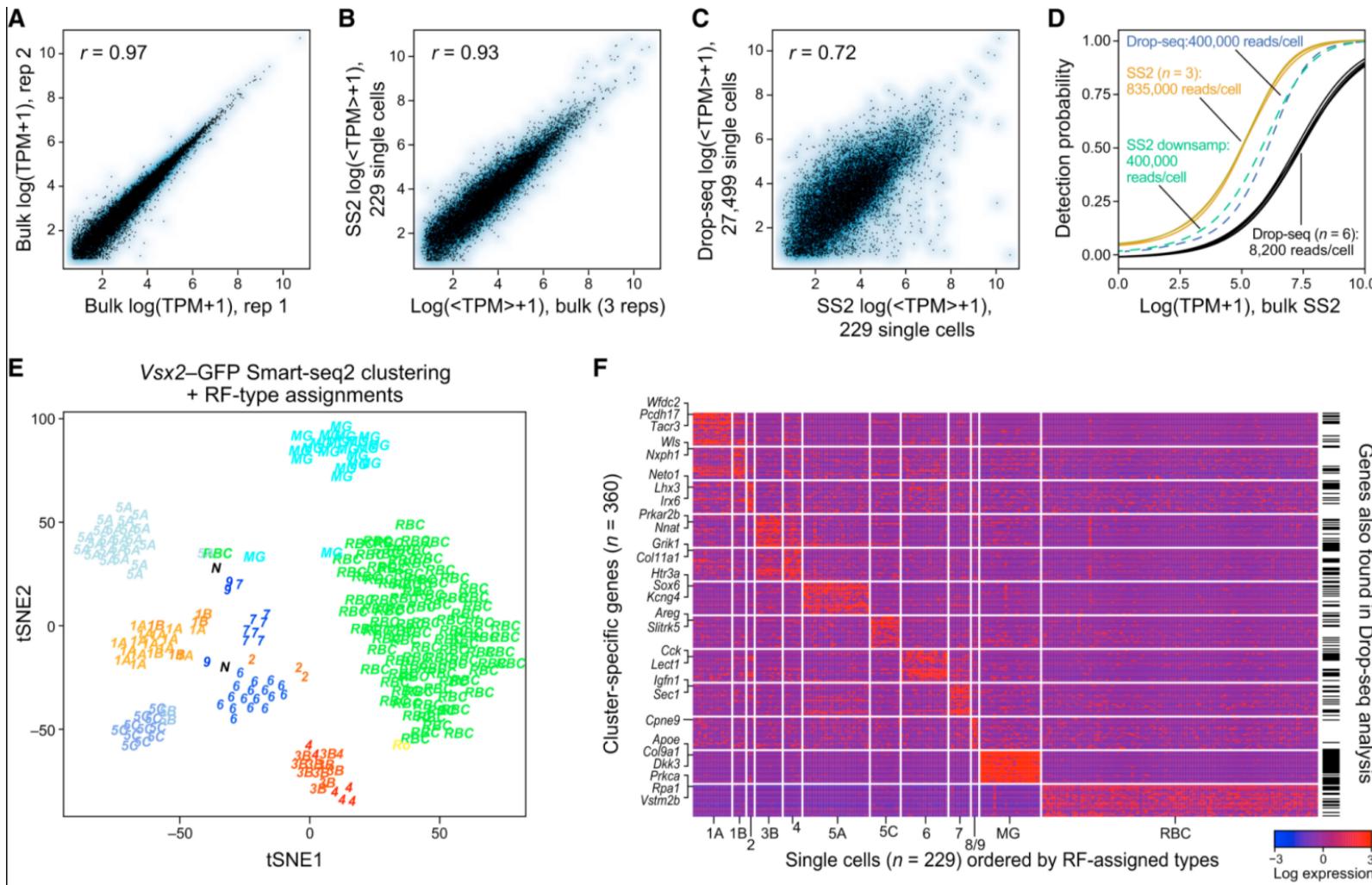
Resource

Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics

Karthik Shekhar ^{1, 9}, Sylvain W. Lapan ^{2, 7, 9}, Irene E. Whitney ^{4, 9}, Nicholas M. Tran ⁴, Evan Z. Macosko ^{2, 5, 6}, Monika Kowalczyk ¹, Xian Adiconis ^{1, 5}, Joshua Z. Levin ^{1, 5}, James Nemesh ^{2, 5, 6}, Melissa Goldman ^{2, 5}, Steven A. McCarroll ^{2, 5, 6}, Constance L. Cepko ^{2, 3, 7}  , Aviv Regev ^{1, 7, 8}  , Joshua R. Sanes ^{4, 10}  



Identifying cell types from scRNA-seq



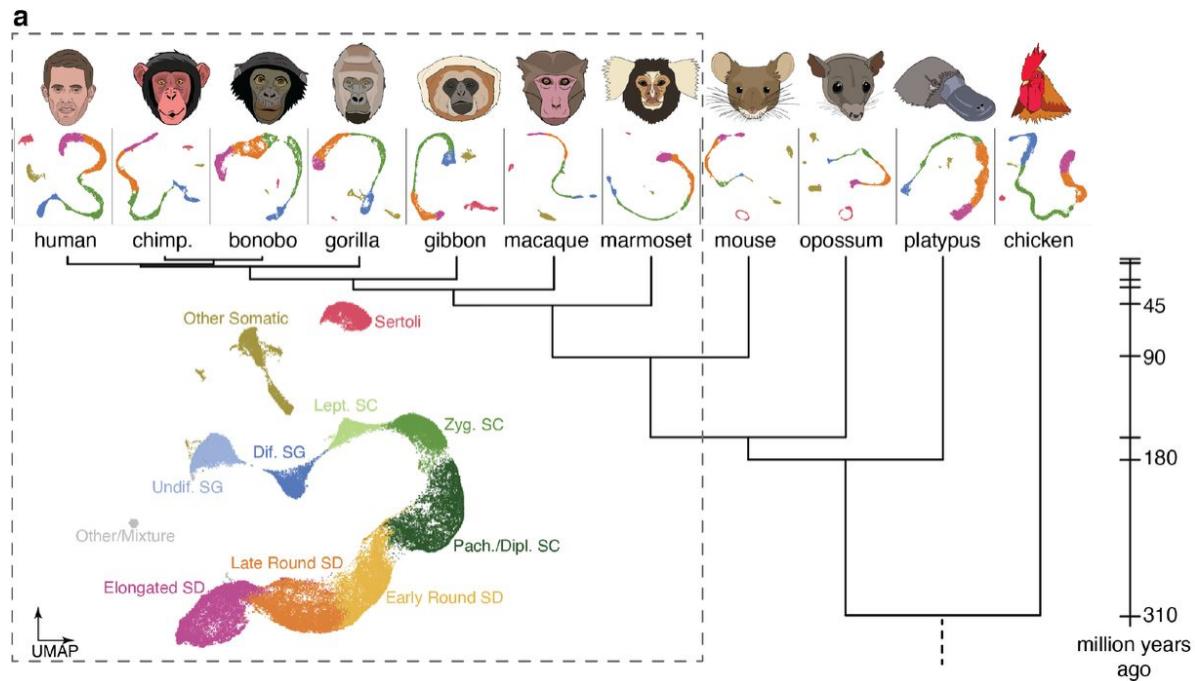
- **SS2 (plate)** experiments tend to be sequenced deeper per cell – **more genes detected**
- **Droplet** experiments have less depth, but **capture more cells**
- **To identify cell types, more cells is better than more genes**

Identifying trajectories from scRNA-seq

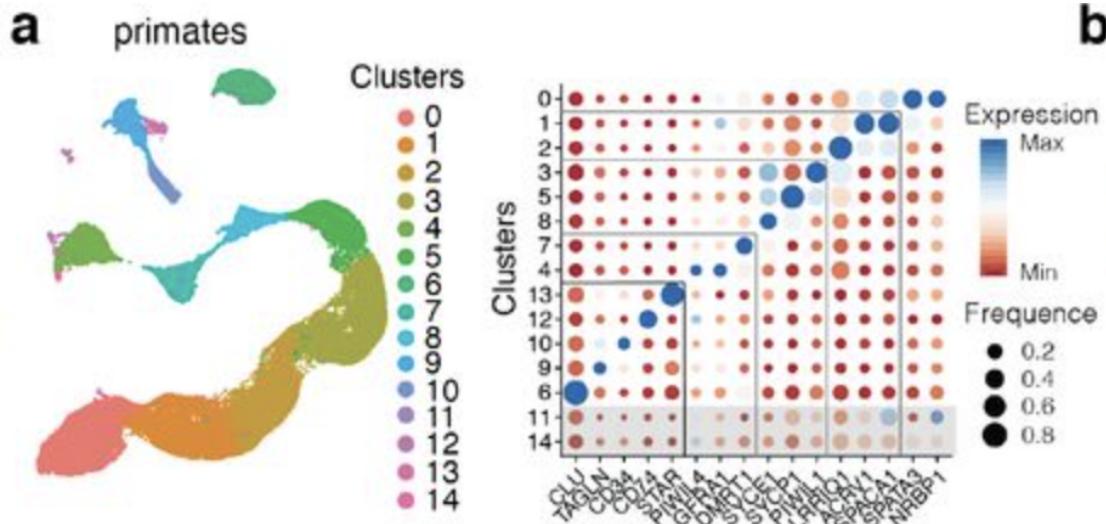
The molecular evolution of spermatogenesis across mammals

✉ Florent Murat, ✉ Noe Mbengue, ✉ Sofia Boeg Winge, ✉ Timo Trefzer, Evgeny Leushkin, ✉ Mari Sepp,
✉ Margarida Cardoso-Moreira, Julia Schmidt, Celine Schneider, Katharina Mößinger, Thoomke Brüning,
✉ Francesco Lamanna, Meritxell Riera Belles, ✉ Christian Conrad, Ivanela Kondova, ✉ Ronald Bontrop,
✉ Rüdiger Behr, ✉ Philipp Khaitovich, Svante Pääbo, ✉ Tomas Marques-Bonet, Frank Grützner,
✉ Kristian Almstrup, ✉ Mikkel Heide Schierup, ✉ Henrik Kaessmann

- scRNA-seq data can also identify **continuous** effects
 - UMAP and tSNE can give an indication that a trajectory may exist, but should not be used to model it
- **cross-species comparisons** can reveal evolutionary history of cell types

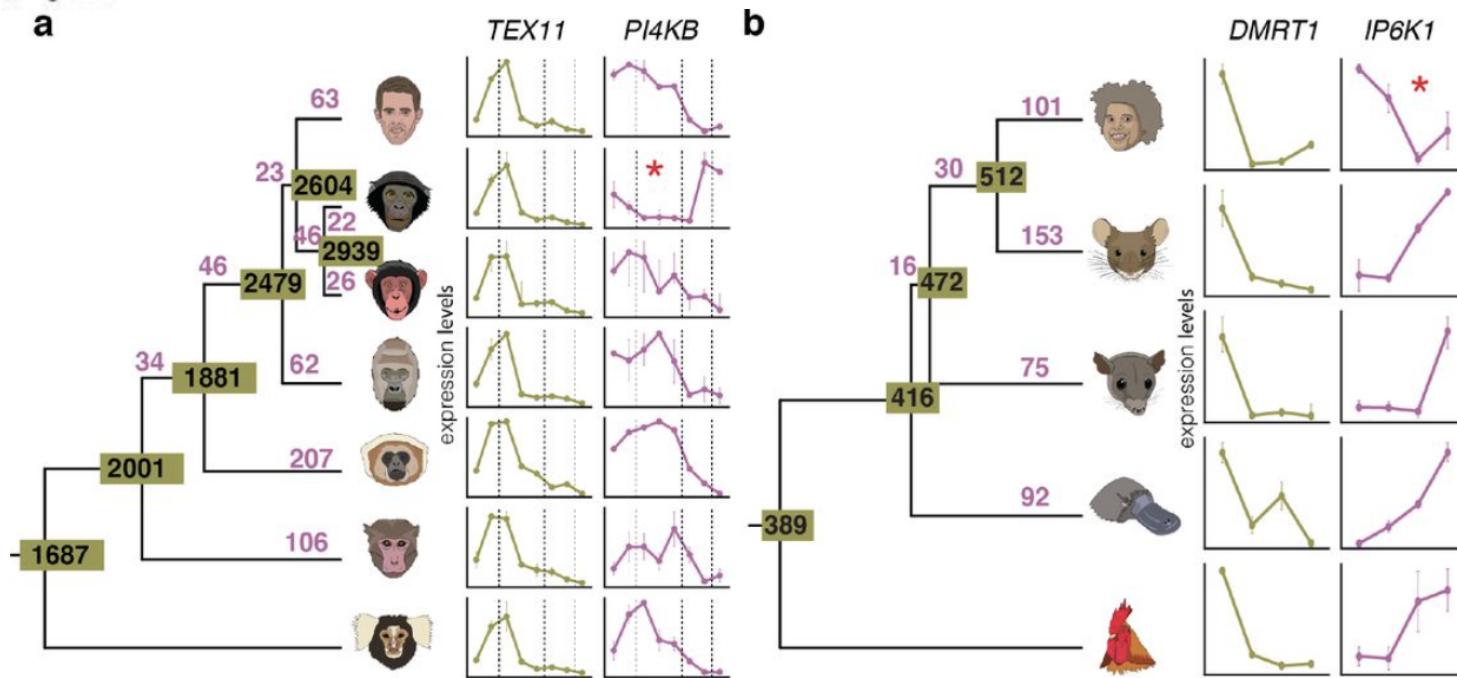


Identifying trajectories from scRNA-seq



b

Pseudotime analysis allows for the identification of **continuously variable genes**



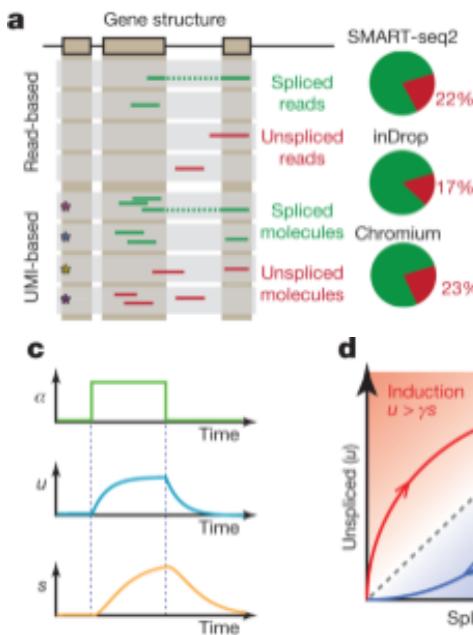
Identifying trajectories from scRNA-seq – RNA velocity

RNA velocity of single cells

Gioele La Manno, Ruslan Soldatov, Amit Zeisel, Emelie Braun, Hannah Hochgerner, Viktor Petukhov, Katja Lidschreiber, Maria E. Kastriti, Peter Lönnerberg, Alessandro Furlan, Jean Fan, Lars E. Borm, Zehua Liu, David van Bruggen, Jimin Guo, Xiaoling He, Roger Barker, Erik Sundström, Gonçalo Castelo-Branco, Patrick Cramer, Igor Adameyko, Sten Linnarsson & Peter V. Kharchenko

Nature 560, 494–498 (2018) | [Cite this article](#)

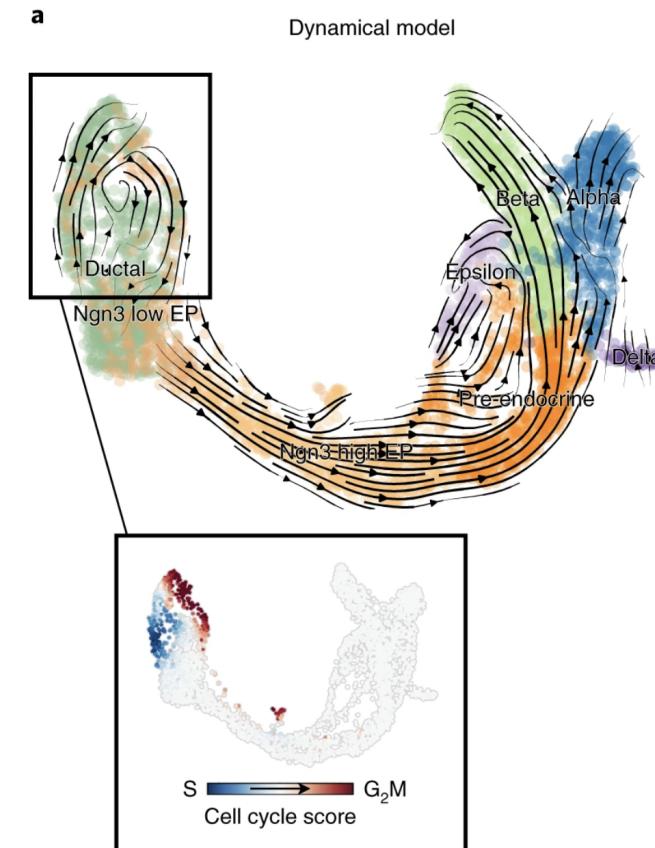
- By incorporating information about **exonic and intronic reads**, we can model a directionality for a cell's trajectory – **RNA velocity**
- **Intronic reads** are obtained from the **nucleus**, so should generally be **more recent** than exonic reads



Generalizing RNA velocity to transient cell states through dynamical modeling

Volker Bergen, Marius Lange, Stefan Peidli, F. Alexander Wolf & Fabian J. Theis

Nature Biotechnology 38, 1408–1414 (2020) | [Cite this article](#)



<https://www.nature.com/articles/s41586-018-0414-6>

<https://www.nature.com/articles/s41587-020-0591-3>

Two major platforms for scRNA-seq analysis



```
library(Seurat)

# load data and create Seurat object
pbmc.data = Read10X(data.dir = "../data/pbmc3k/filtered_gene_bc_matrices/hg19/")
pbmc = CreateSeuratObject(counts = pbmc.data, min.cells = 3, min.features = 200)

# normalise data
pbmc = NormalizeData(pbmc, normalization.method = "LogNormalize", scale.factor =
10000)

# find highly variable genes
pbmc = FindVariableFeatures(pbmc, selection.method = "vst", nfeatures = 2000)

# scale data
pbmc = ScaleData(pbmc, features = rownames(pbmc))

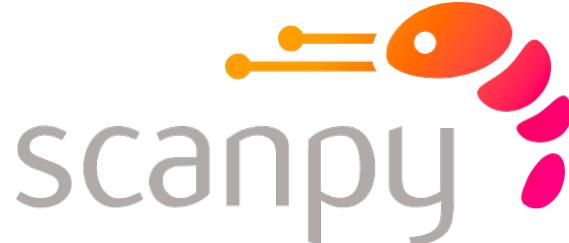
# linear dimensionality reduction
pbmc = RunPCA(pbmc, features = VariableFeatures(object = pbmc))

# clustering
pbmc = FindNeighbors(pbmc, dims = 1:40)
pbmc = FindClusters(pbmc, resolution = 0.5)

# non-linear dimensionality reduction
pbmc = RunUMAP(pbmc, dims = 1:10)

# find marker genes
pbmc.mk = FindAllMarkers(pbmc, only.pos = TRUE, min.pct = 0.25, logfc.threshold =
0.25)
```

<https://satijalab.org/seurat/>



```
import scanpy as sc

# load data and create AnnData (scanpy) object
adata = sc.read_10x_mtx('data/filtered_gene_bc_matrices/hg19/',
var_names='gene_symbols')

# normalise data
sc.pp.normalize_total(adata, target_sum=1e4)
sc.pp.log1p(adata)

# find highly variable genes
sc.pp.highly_variable_genes(adata, min_mean=0.0125, max_mean=3,
min_disp=0.5)

# scale data
sc.pp.scale(adata, max_value=10)

# linear dimensionality reduction
sc.tl.pca(adata, svd_solver='arpack')

# clustering
sc.pp.neighbors(adata, n_neighbors=10, n_pcs=40)
sc.tl.leiden(adata)

# non-linear dimensionality reduction
sc.tl.umap(adata)

# find marker genes
sc.tl.rank_genes_groups(adata, 'leiden', method='t-test')
```

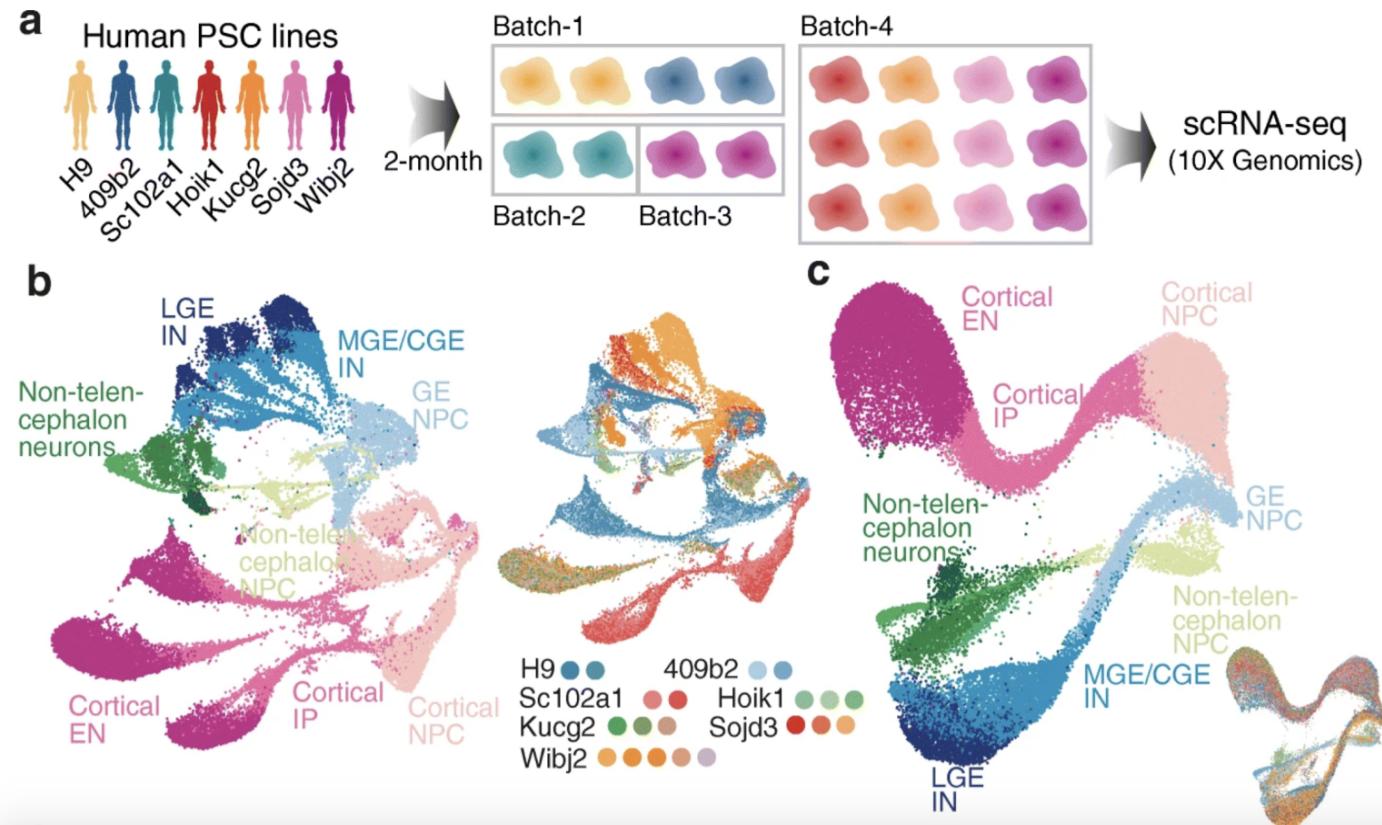
<https://scanpy.readthedocs.io/en/stable/> <https://www.sc-best-practices.org/preamble.html>

Topics:

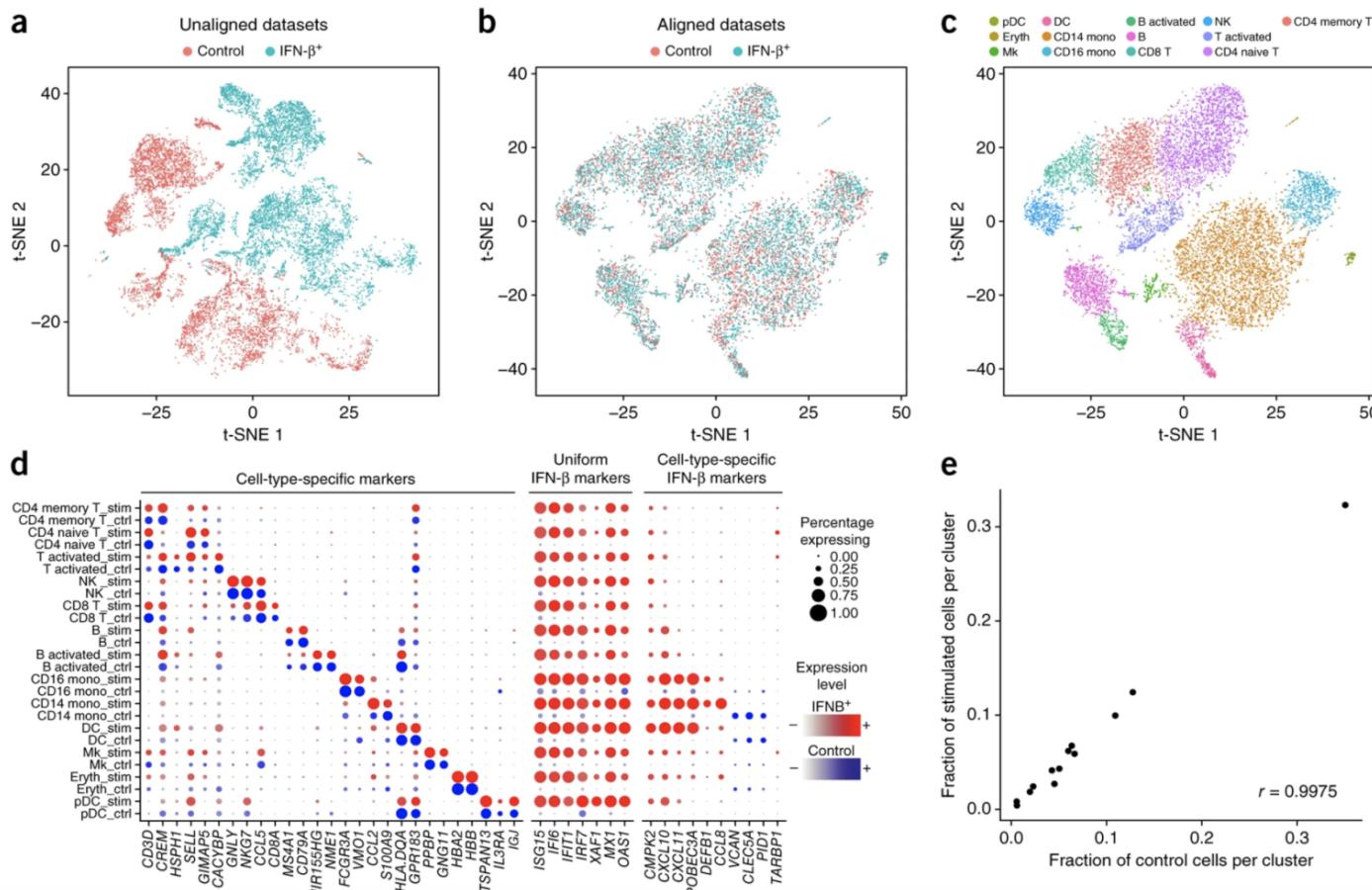
- The emergence of single-cell sequencing
- How to get mRNA from single-cells
- Processing and analysis steps
- **Data integration and multi-modal data**
- Areas of intense, active development

Data integration – correct unwanted variation

- Gene expression matrices reflect a variety of factors, both **technical and biological**
- The usual goal with scRNA-seq is to **represent cell types**
- Computational integration tools work on a low-dimension space (e.g. PCA) to **correct for unwanted technical variation** (e.g. donor, experiment day, isolation protocol, or lab)



Correct wanted (?) variation

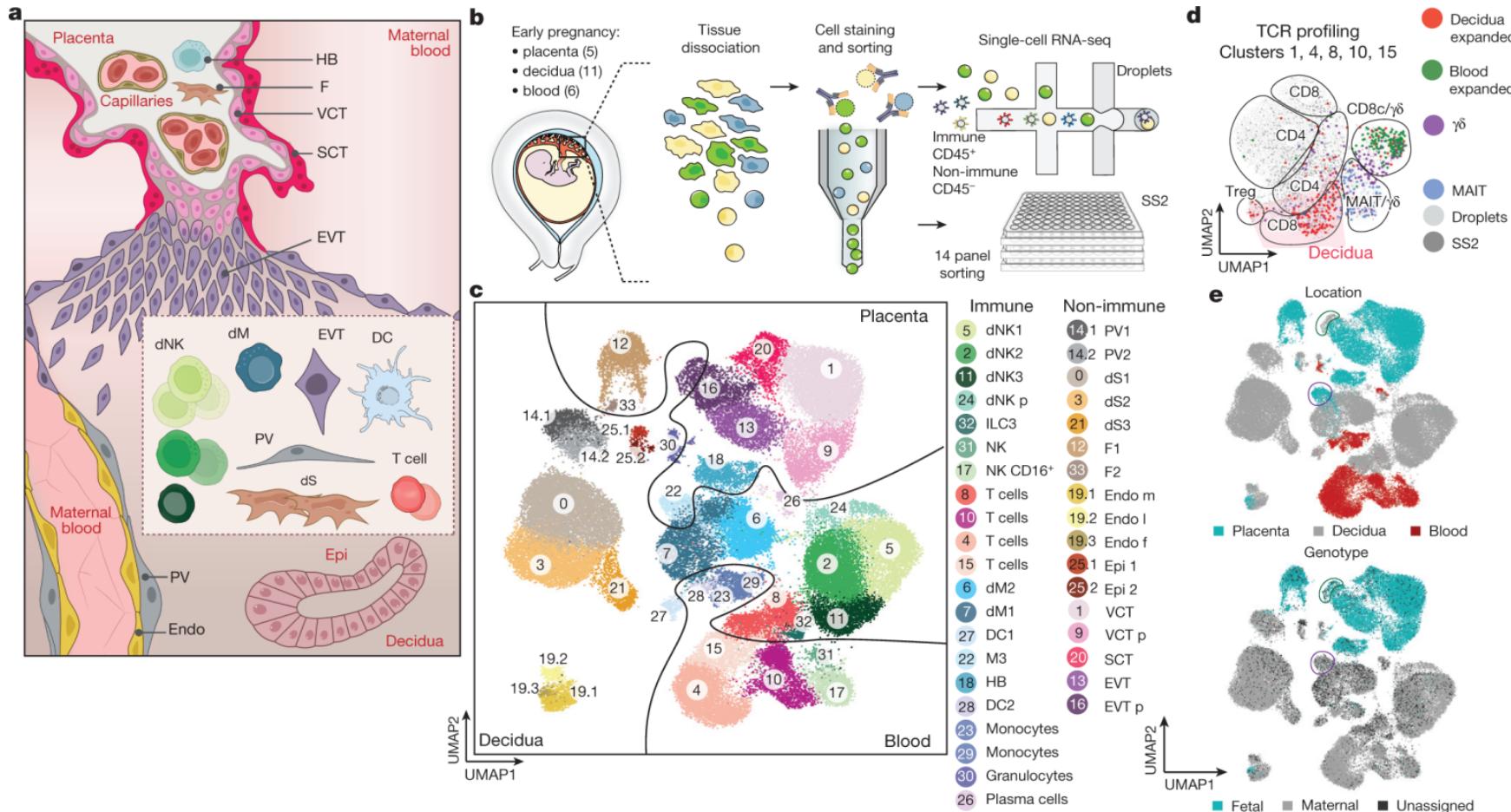


It can also happen that **cell types are harder to identify in multifactorial experiments**, due to low biological variability

Similarly, some **lowly-represented cell types** only appear once the data is integrated

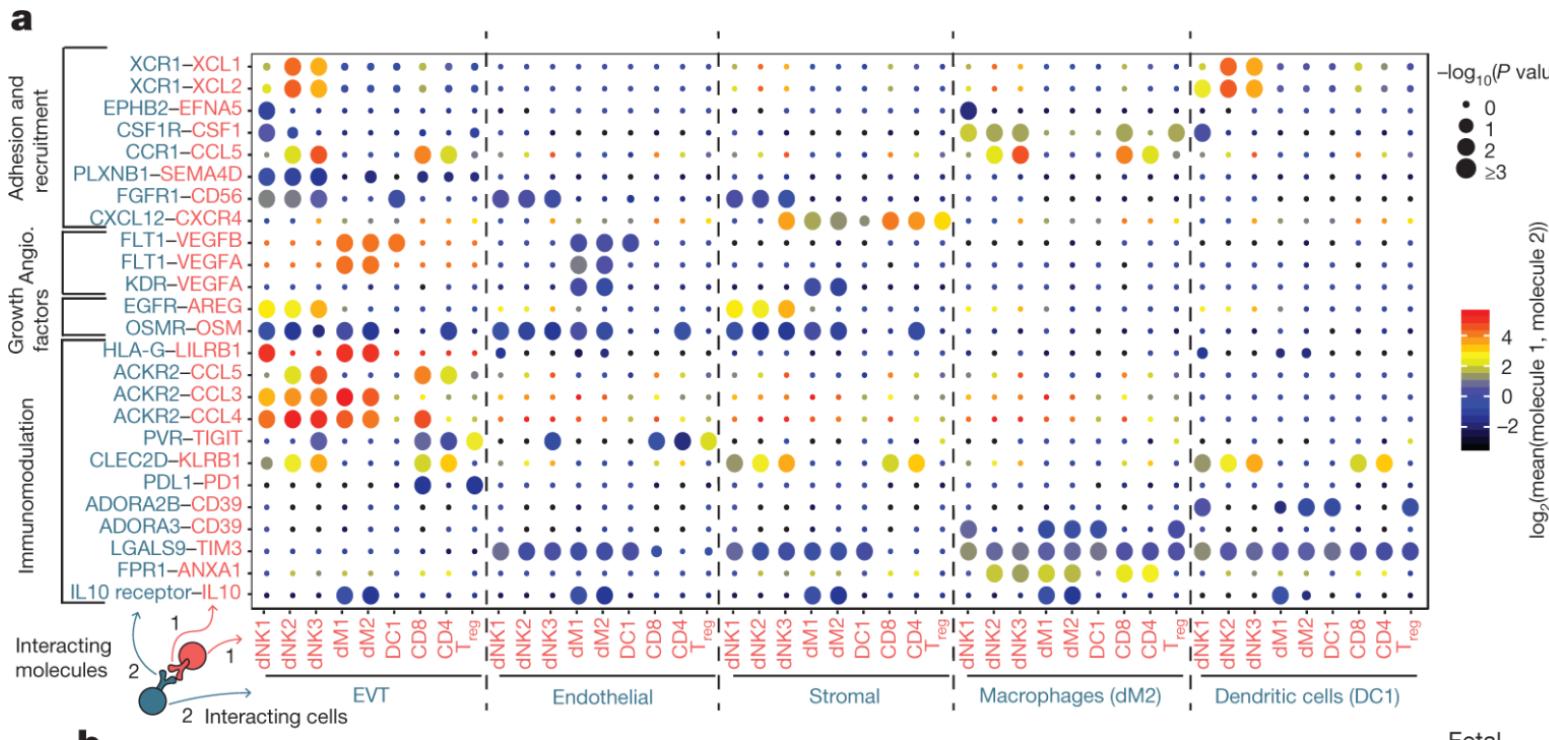
Common biological factors that are corrected for are tissue, treatment, or species

The importance of using multiple types of data

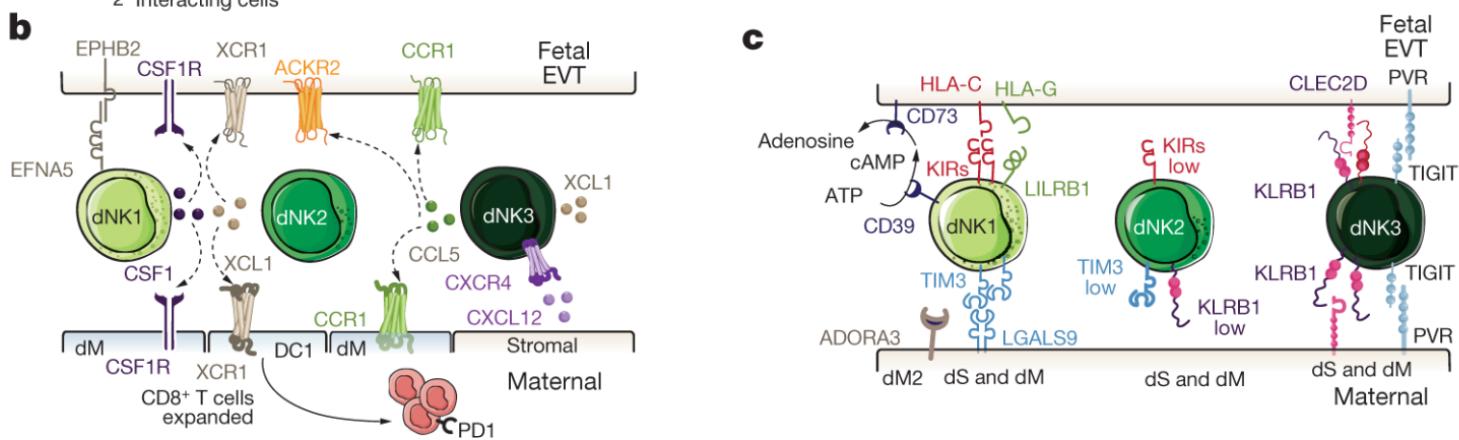


- sequencing reads (with the help of a reference genome) can help call variants and **assign cell origin to different individuals**
- this can be used in **multiplexed experiments**, but is useful here to distinguish maternal and fetal cells

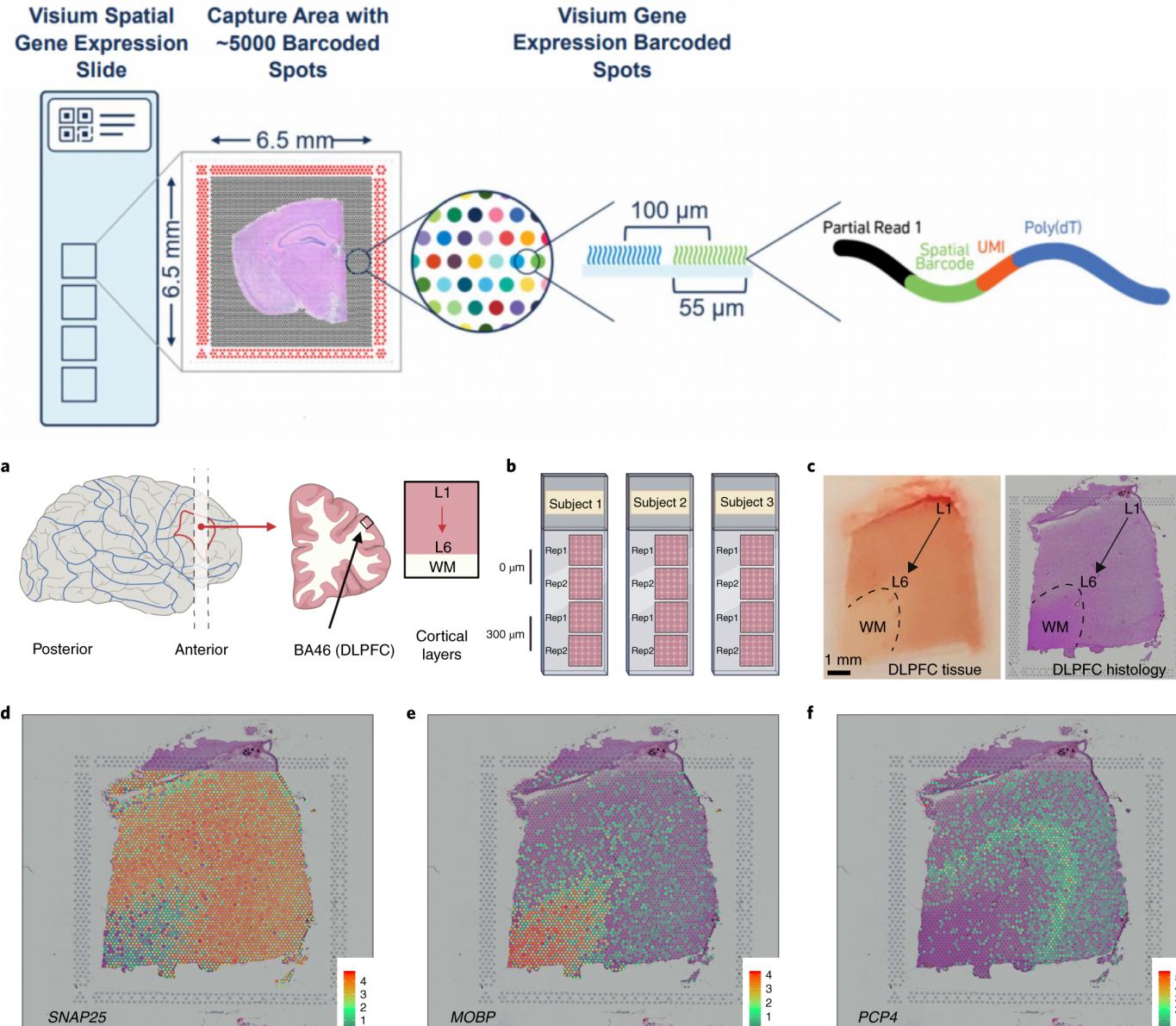
The importance of using multiple types of data



scRNA-seq can be used to infer cell-cell communication events

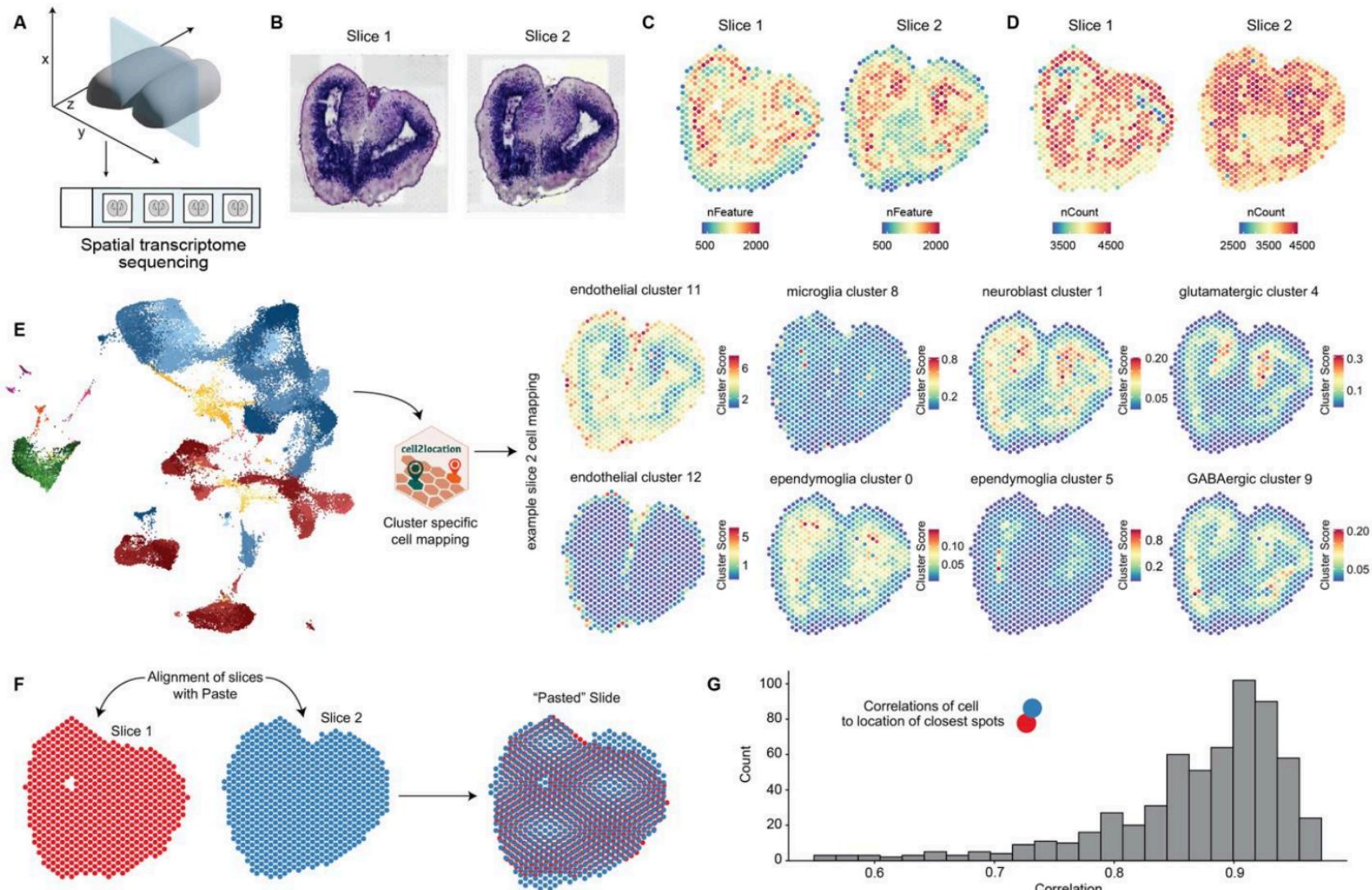


Spatial transcriptomics



- Spatial transcriptomics brings together histology and gene expression
- Many technologies with different throughputs (few to all genes) and resolution (up to subcellular)

Spatial transcriptomics



scRNA-seq can be mapped spatially

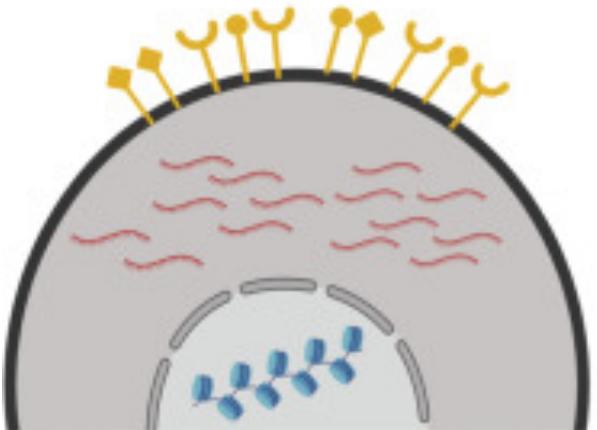
Fig. S5. Overview of axolotl telencephalon spatial transcriptomics.

Topics:

- The emergence of single-cell sequencing
- How to get mRNA from single-cells
- Processing and analysis steps
- Data integration and multi-modal data
- Areas of intense, active development

Beyond RNA

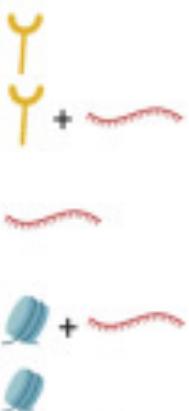
(A)



Molecule	Protocol
Protein	CITE-seq SCoPe-MS
RNA	scRNA-seq VDJ-seq snRNA-seq
DNA/ chromatin	scATAC-seq CUT&Tag scDNA-seq

(B)

- Multiplex samples and donors
- Preselected surface proteins in novel subpopulations
- Characterize new immune cell types and states
(TCR/BCR)-specific expression
- Map regulatory networks controlling immune response
- Reveal antigen variation in pathogens



Trends in Immunology

We can recover more than just RNA from single-cells:

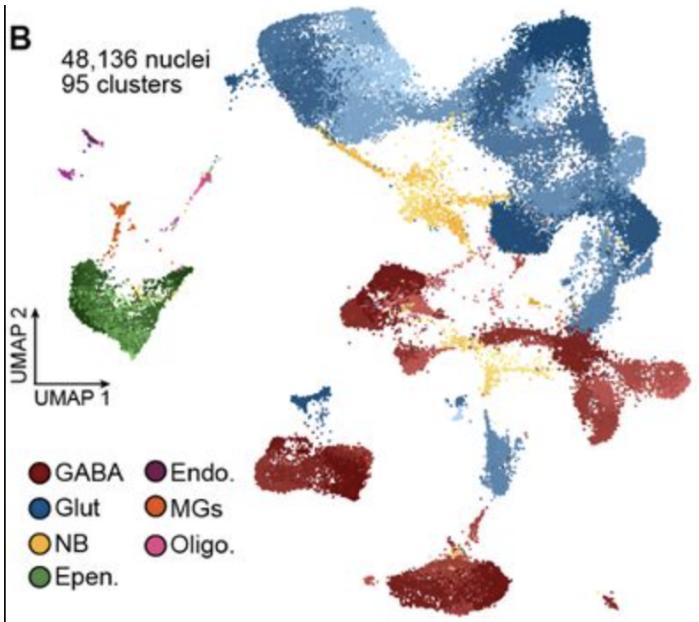
- targeted surface-proteins (CITE-seq)
- TCR/BCR (variable immune transcripts)
- open chromatin (ATAC-seq)
- Histone modifications (CUT&Tag)
- methylation
- proteins (using mass spec)

Single-cell multiomics

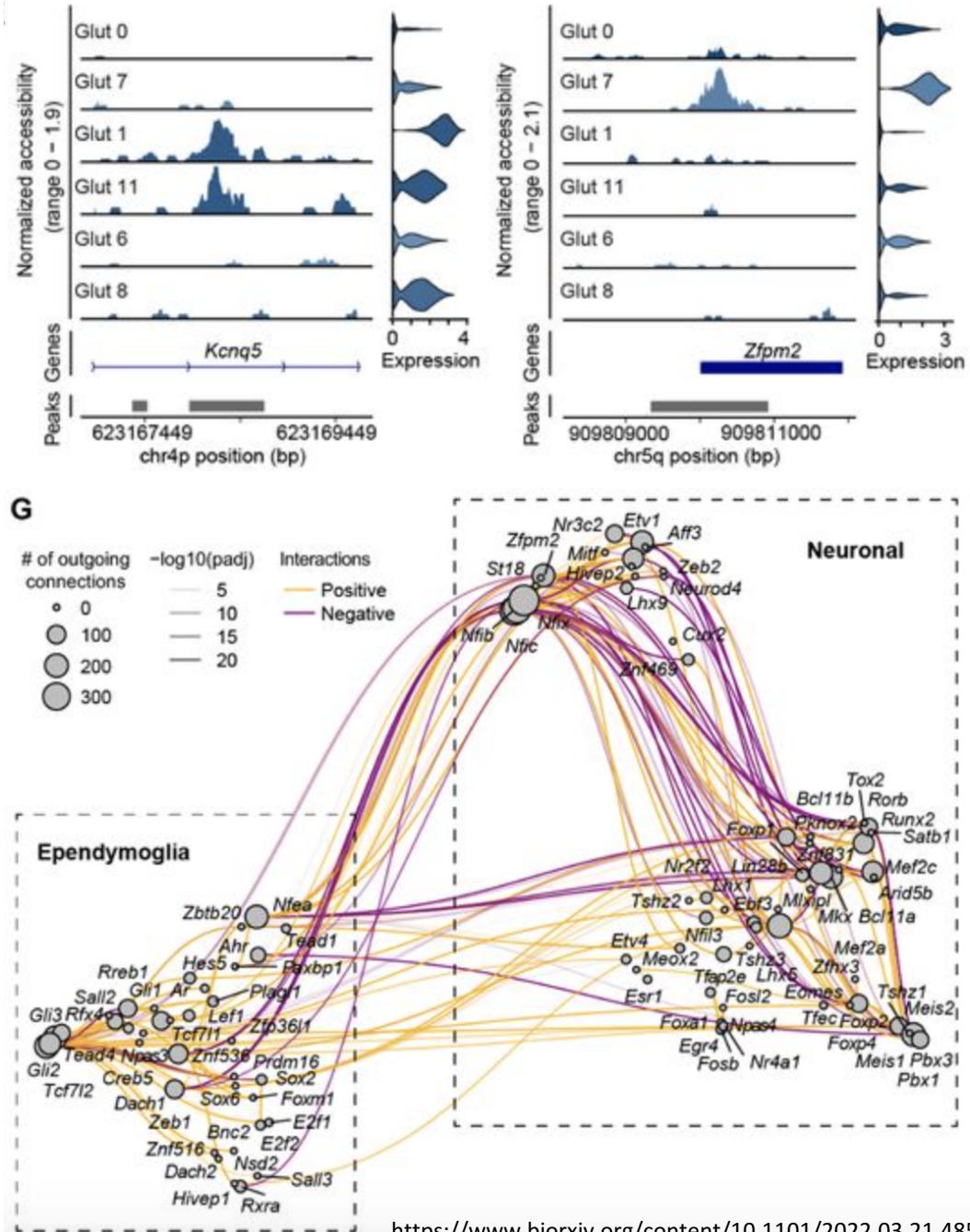
Single-cell analyses of axolotl forebrain organization, neurogenesis, and regeneration

Katharina Lust, Ashley Maynard, Tomás Gomes, Jonas Simon Fleck, J. Gray Camp, Elly M. Tanaka, Barbara Treutlein

doi: <https://doi.org/10.1101/2022.03.21.485045>



Obtaining RNA + open chromatin from the same cell allows for the reconstruction of regulatory networks



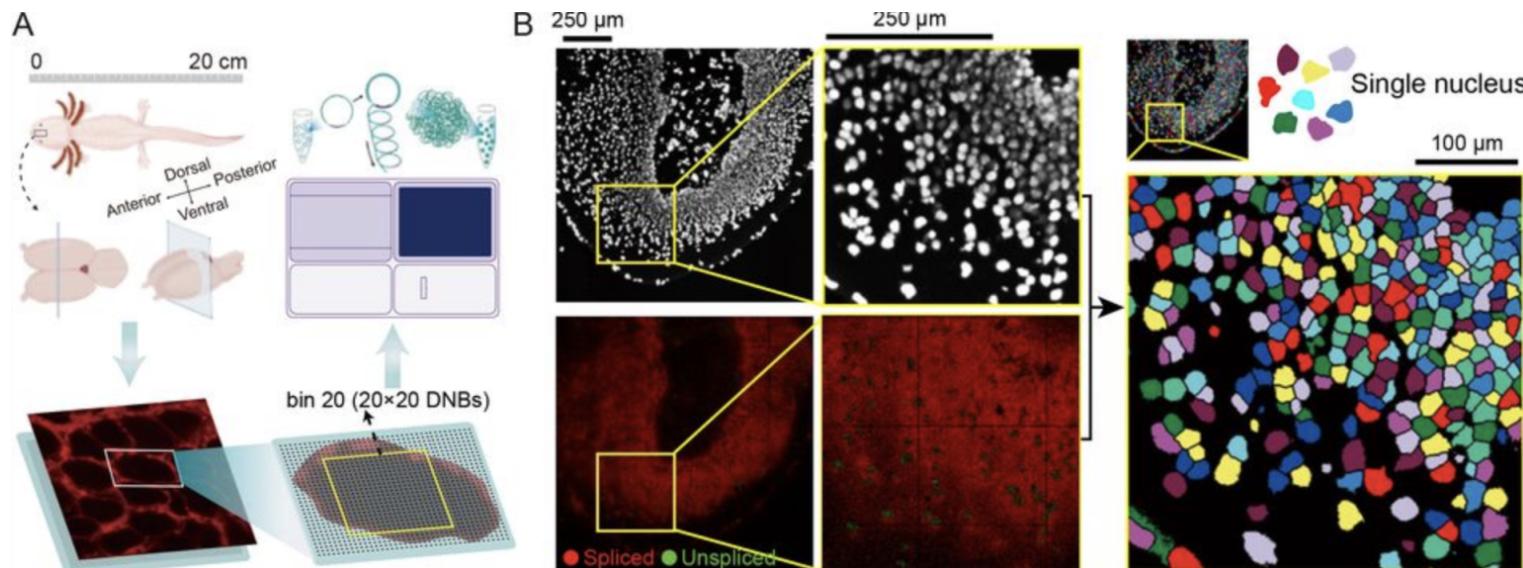
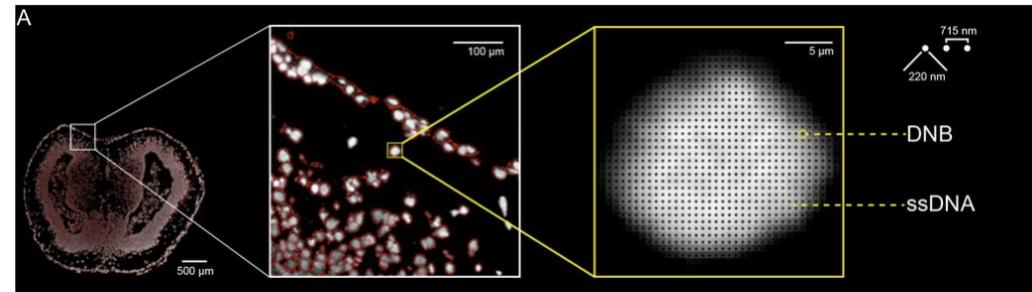
<https://www.biorxiv.org/content/10.1101/2022.03.21.485045v1.full>

Single cells... in space?

Spatiotemporal transcriptome at single-cell resolution reveals key radial glial cell population in axolotl telencephalon development and regeneration

Xiaoyu Wei, Sulei Fu, Hanbo Li, Yang Liu, Shuai Wang, Weimin Feng, Yunzhi Yang, Xiawei Liu, Yan-Yun Zeng, Mengnan Cheng, Yiwei Lai, Xiaojie Qiu, Liang Wu, Nannan Zhang, Yujia Jiang, Jiangshan Xu, Xiaoshan Su, Cheng Peng, Lei Han, Wilson Pak-Kin Lou, Chuanyu Liu, Yue Yuan, Kailong Ma, Tao Yang,  Xiangyu Pan, Shang Gao, Ao Chen, Miguel A. Esteban, Huanming Yang, Jian Wang, Guangyi Fan, Longqi Liu, Liang Chen, Xun Xu, Ji-Feng Fei,  Ying Gu

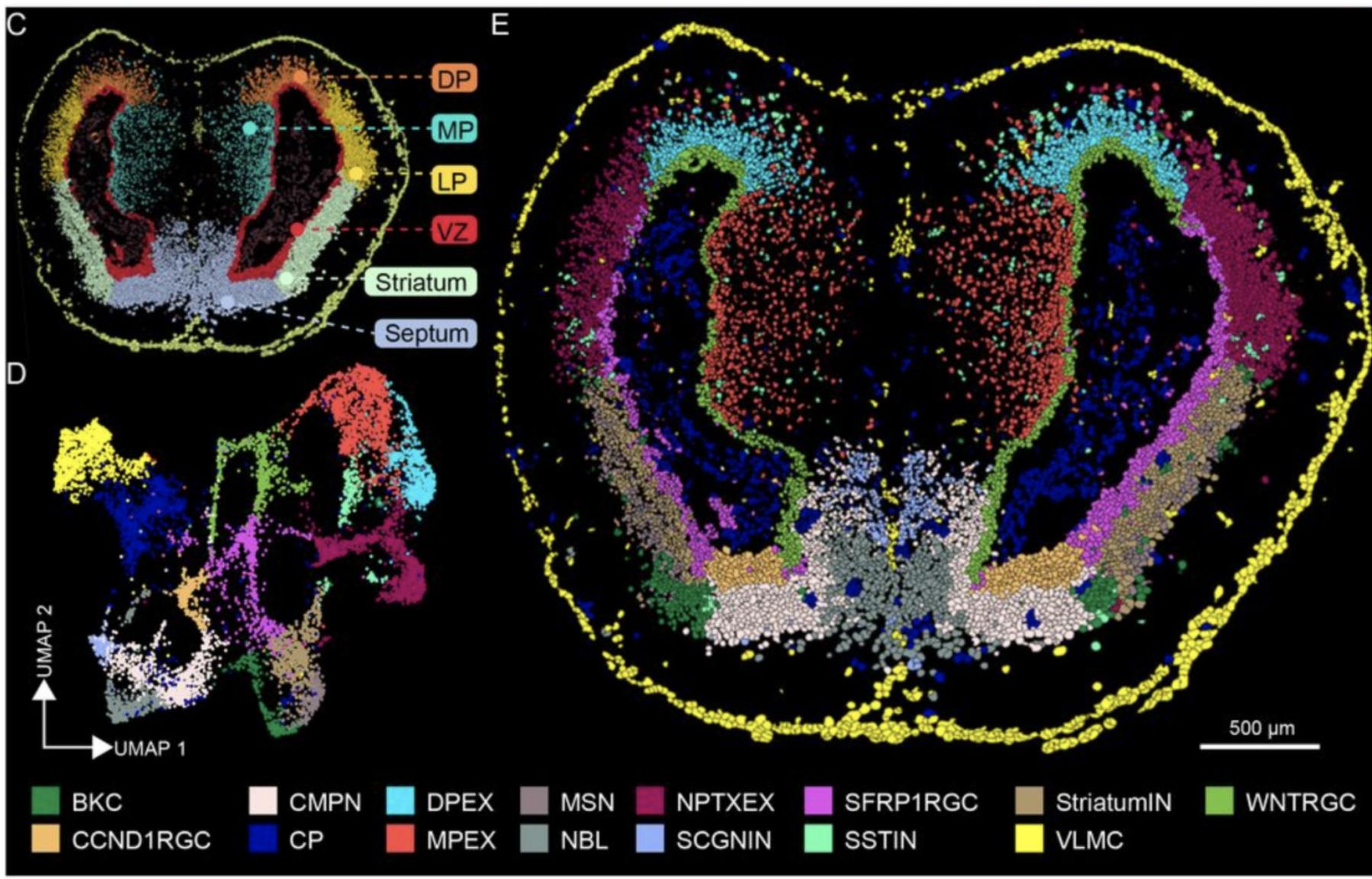
doi: <https://doi.org/10.1101/2021.10.23.465550>



Spatial information is usually destroyed during single-cell isolation

Recent methods are now trying to maintain or approximate it

Single cells... in space?



Deep learning for single-cell

A Python library for probabilistic analysis of single-cell omics data

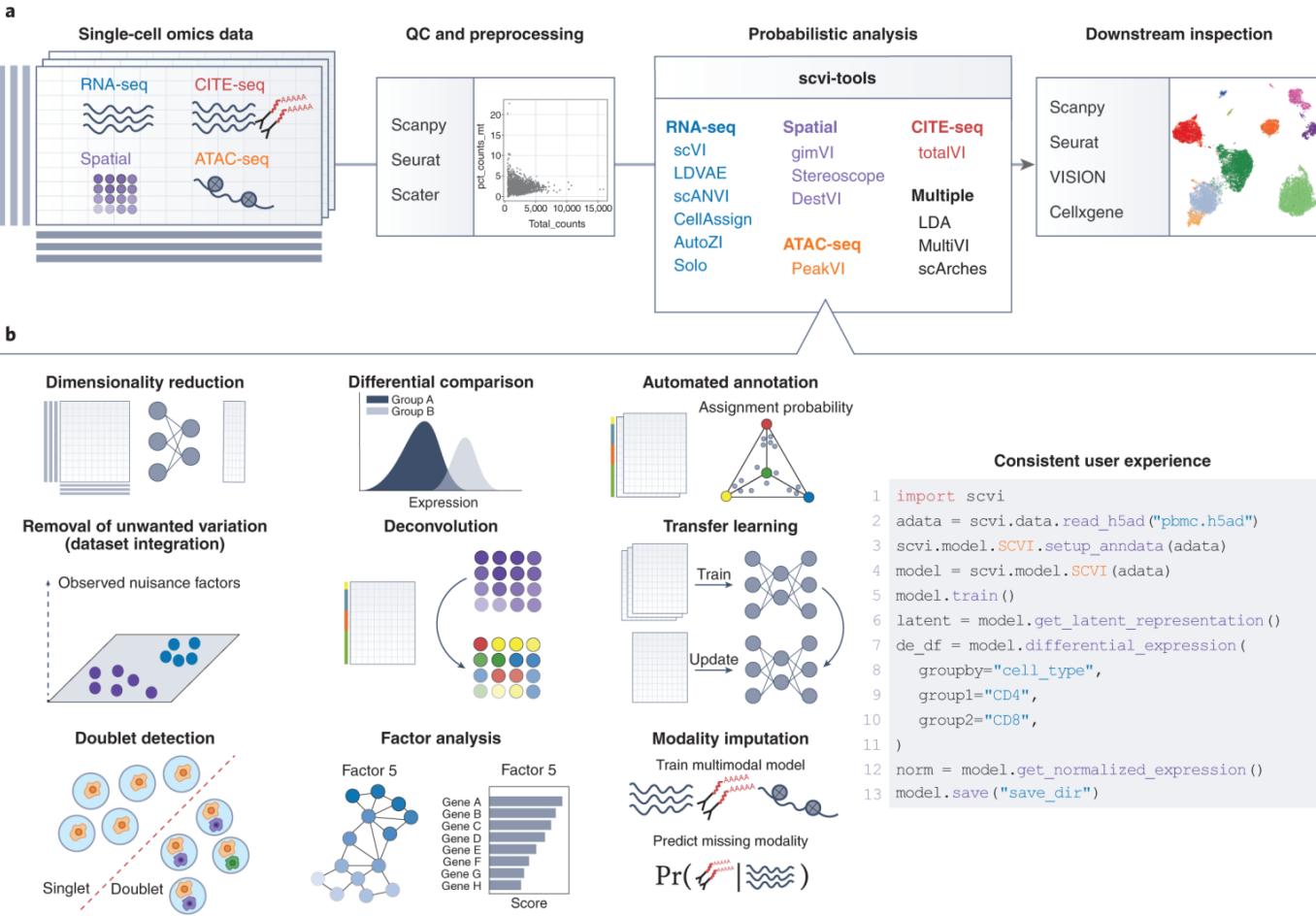
Adam Gayoso, Romain Lopez, Galen Xing, Pierre Boyeau, Valeh Valiollah Pour Amiri, Justin Hong, Katherine Wu, Michael Jayasuriya, Edouard Mehlman, Maxime Langevin, Yining Liu, Jules Samaran, Gabriel Misrachi, Achille Nazaret, Oscar Clivio, Chenling Xu, Tal Ashuach, Mariano Gabitto, Mohammad Lotfollahi, Valentine Svensson, Eduardo da Veiga Beltrame, Vitalii Kleshchevnikov, Carlos Talavera-López, Lior Pachter, Fabian J. Theis, Aaron Streets, Michael I. Jordan, Jeffrey Regier & Nir Yosef 

— Show fewer authors

Nature Biotechnology 40, 163–166 (2022) | [Cite this article](#)

Deep learning strategies for single-cell

- replace all standard analysis
- transfer learning
- Treatment prediction
- integration
- spatial deconvolution
- ...



Scalable, especially when based on GPU

Topics:

- The emergence of single-cell sequencing

scRNA-seq allows for the study of cell biology at whole-transcriptome level

- How to get mRNA from single-cells

Larger datasets allow identification of rarer cell types

- Processing and analysis steps

Data analysis is centred on normalization and clustering

- Data integration and multi-modal data

Single-cell data can be enhanced by exploring different aspects of it (cell-cell communication, splicing, spatial)

- Areas of intense, active development

Different 'omic modalities and different parallel modalities from the same cell

Increased resolution in spatial transcriptomics

Thank you!

Questions?

Tomás Gomes, PhD
 @tomsgoms
tomas.gomes@medicina.ulisboa.pt