2/20/2019

Летняя школа по биоинформатике 2018

Single cell RNA-sequencing: advantages and limitations

Konstantin Okonechnikov
Hopp-Children's Cancer Center at the NCT Heidelberg (KiTZ)
German Cancer Research Center (DKFZ)

Short prefix: use scientific language

- Manuscript standard
- Get used to it:
  - Next Generation Sequencing = Секвенирование нового поколения ~ НГС
  - Paired-end reads = парные прочтения ~ риды
  - Single cell = Одиночная (не замужем?) клетка
  - …
  - ∞
- **Presentation slides are in English**

Single cell RNA-seq: welcome to the future

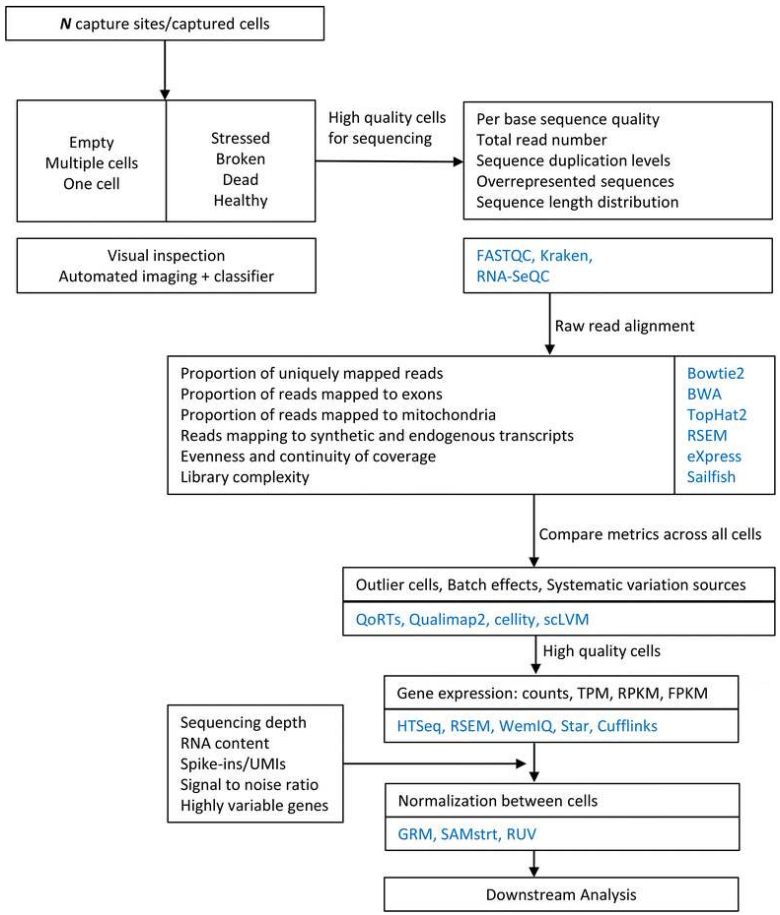
- Detection of cell activity
- Understanding complex cell systems e.g. brain

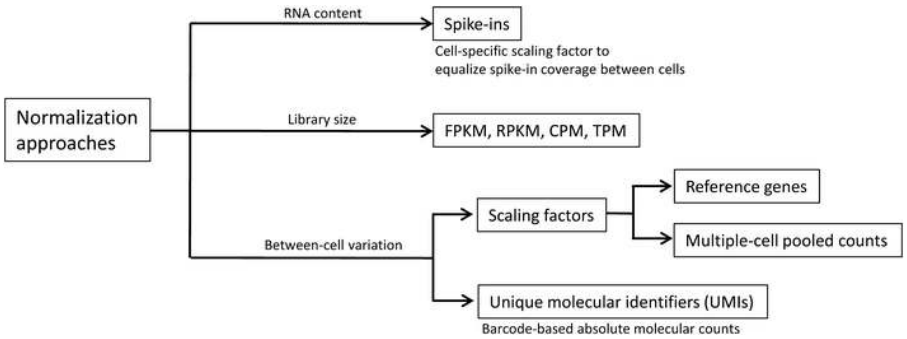
Kageyama et al Front Neurosci. 2018; 12: 315.

A. Rosenberg et al, Science 2018

Single cell RNA-seq: data processing

- General processing highly depends on the **protocol**:
  - No standard procedure





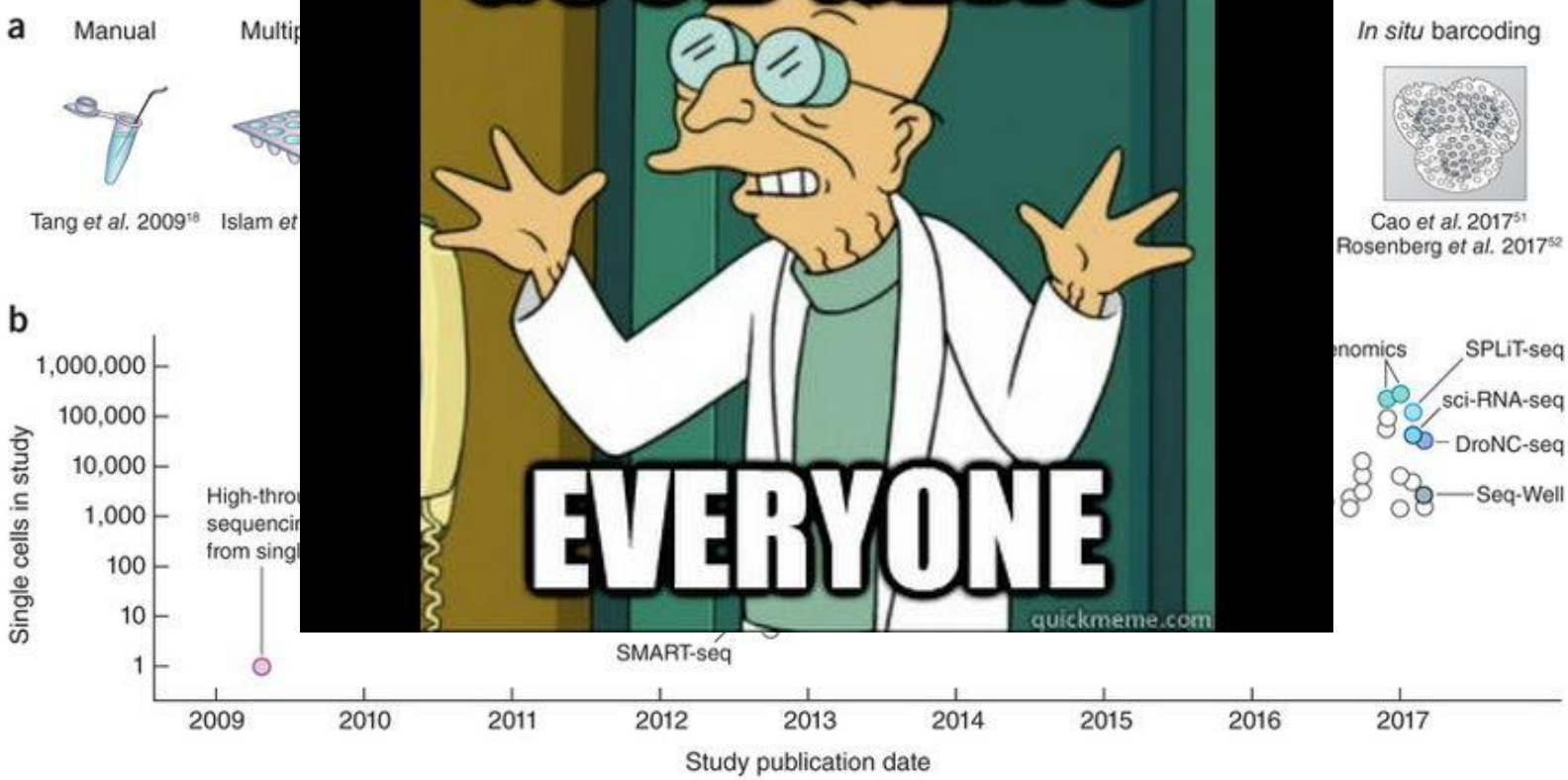
Main novel area:

downstream analysis

Duran, R. et al (2017). Clinical and translational medicine, 6(1), 20.

Single cell RNA-seq: exponential growth

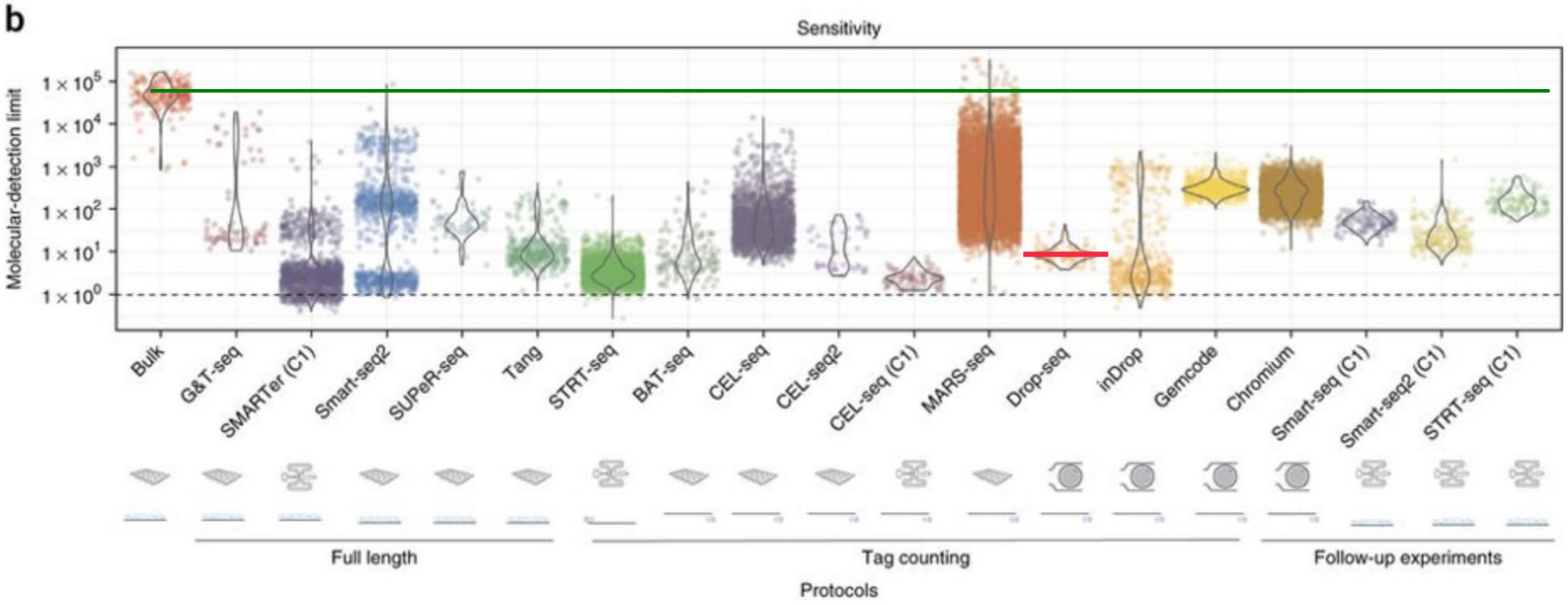
- Various protocols focus on **various targets**
  - Full transcript: ~1K cells ; 3' only: up to 1 000 000 cells
- A lot a



V. Svensson et al Nature Protocols volume 13, pages 599–604 (2018)

Single cell RNA-seq: missing aspects

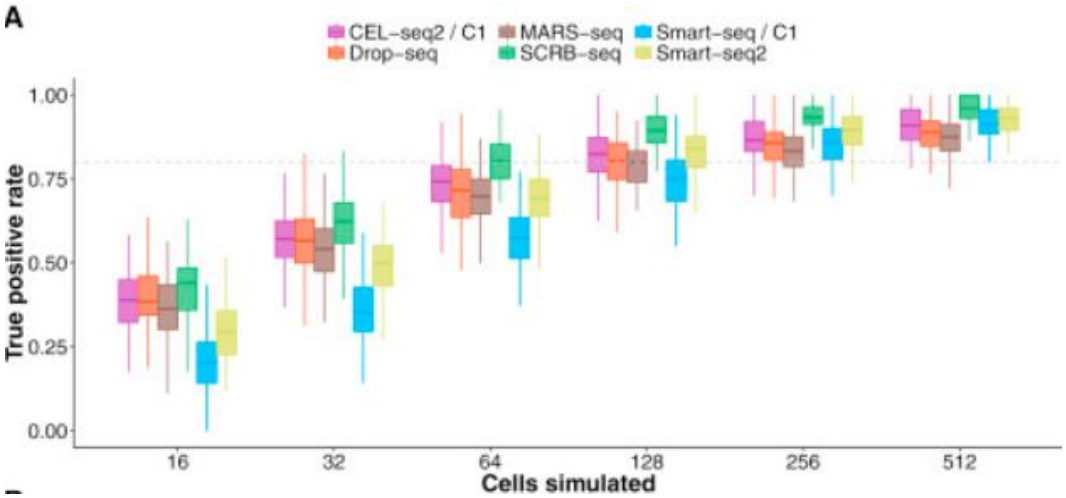
- **Low proportion** of covered genes from cell
- Transcript **subpart** limited e.g. 3' only
- Not working for "**old**" samples e.g. post-mortem

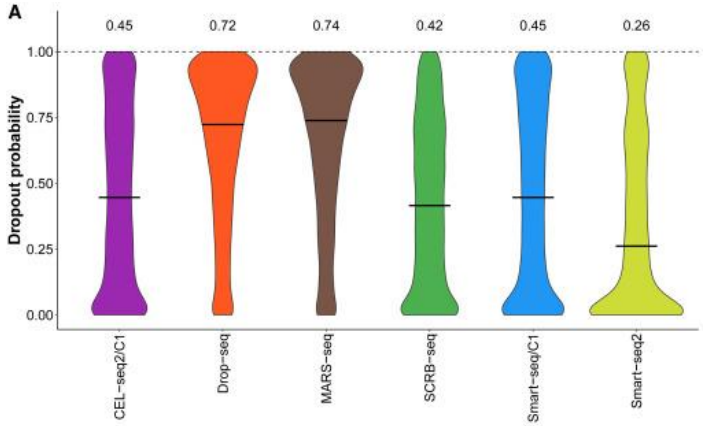


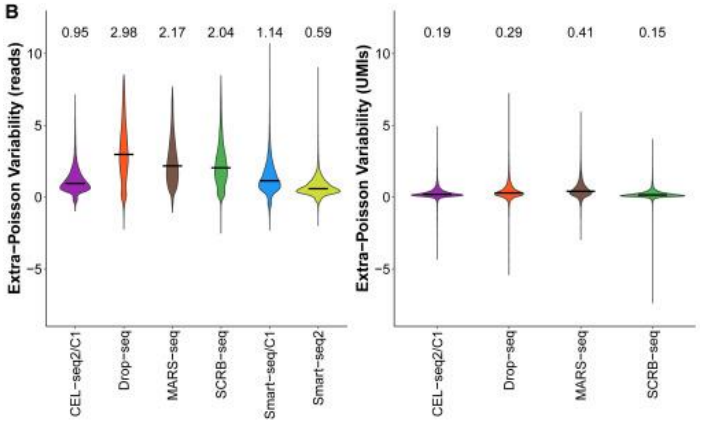
V. Svensson et al Nature Methods volume 14, pages 381–387 (2017)

Global improvements

- Increase in cell counts is highly beneficial
- Accuracy (via **External RNA Control Consortium**)
- Variability: **Unique Molecule Identifiers** benefit





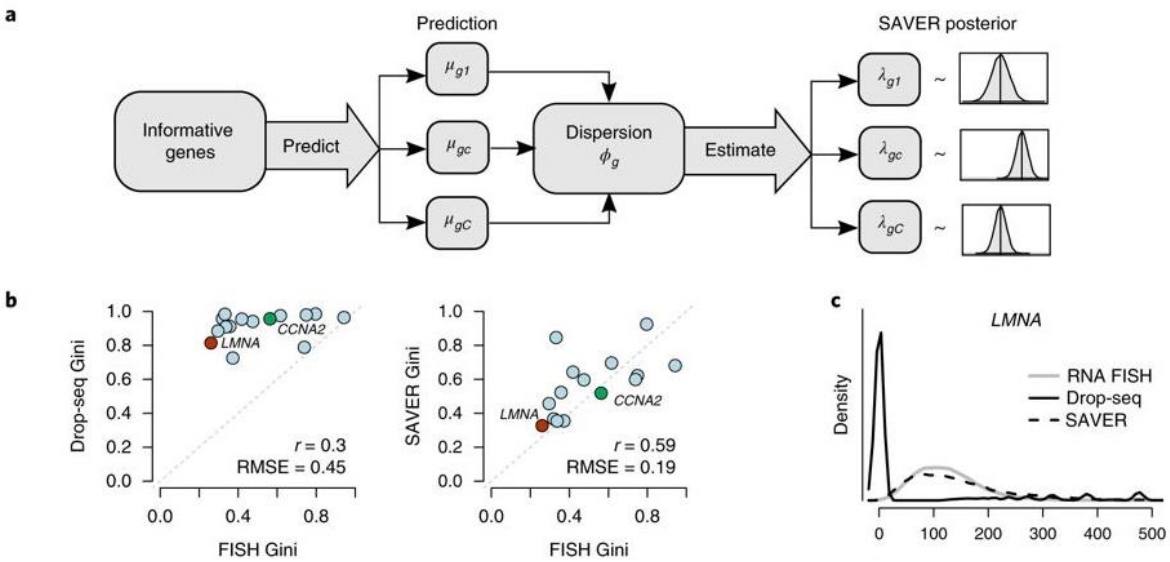


Test on differentially expressed genes through simulation

C. Ziegenhain et al Molecular Cell V65, I4, 2017, Pages 631-643.e4

Predicting genes presence

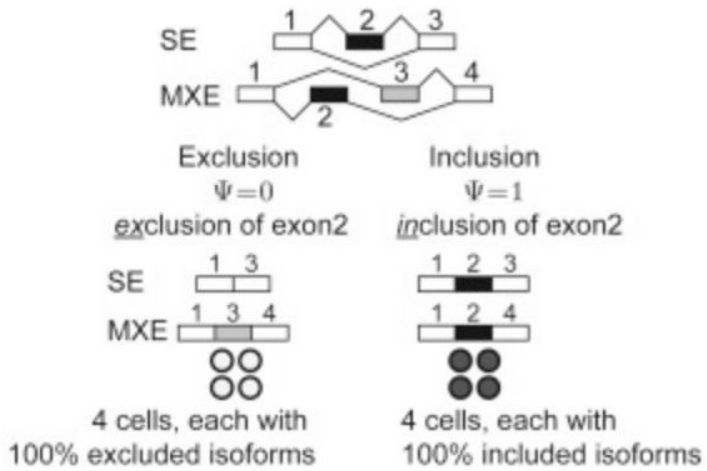
- Computational approach: **gene imputation**
- Use distribution analysis and/or machine learning to understand missing cells patterns
- Several methods available (MAGIC, scImpute, ..)
  - SAVER: gene-to-gene correlations, validated via subsampling and FISH

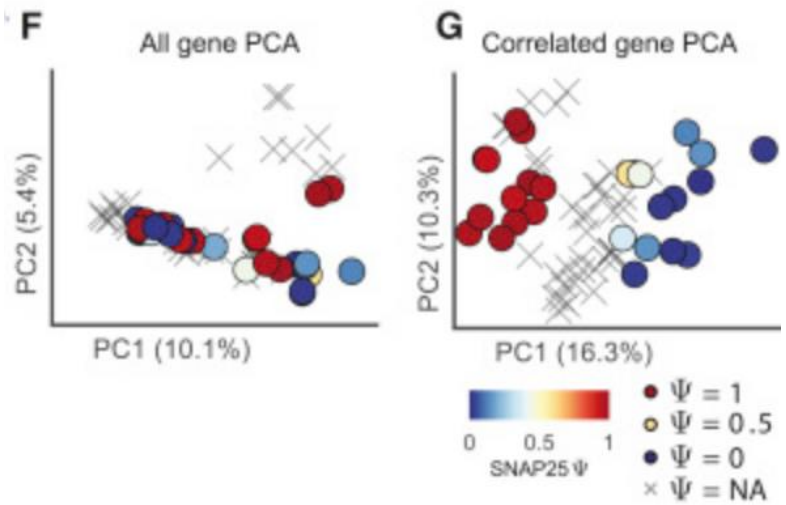


M. Huang et al Nature Methodsvolume 15, pages539–542 (2018)

Full read length protocols

- Cover **whole transcripts** e.g. Smart-seq C1
- Apply ERCC and UMI controls
- Distribution analysis to predict correct evidence of transcripts
  - Bayesian regression, …
  - More precise verification of cell types

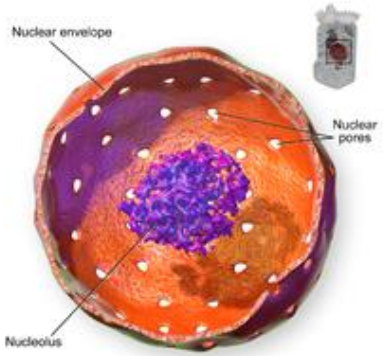




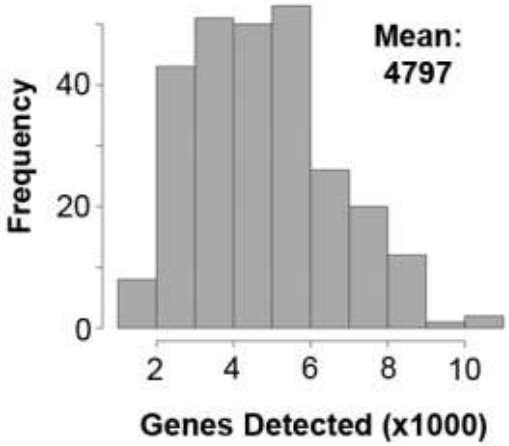
Y. Song et al Molecular Cell, June 29, 2017

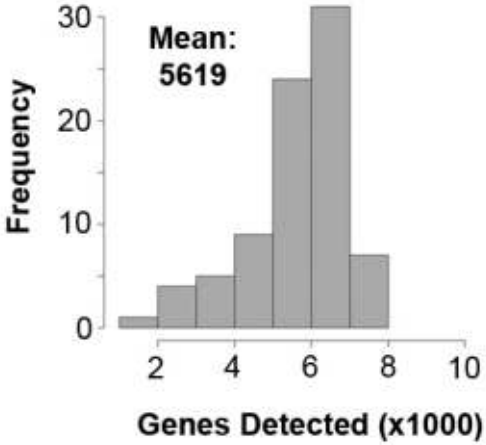
Single nucleus RNA-seq

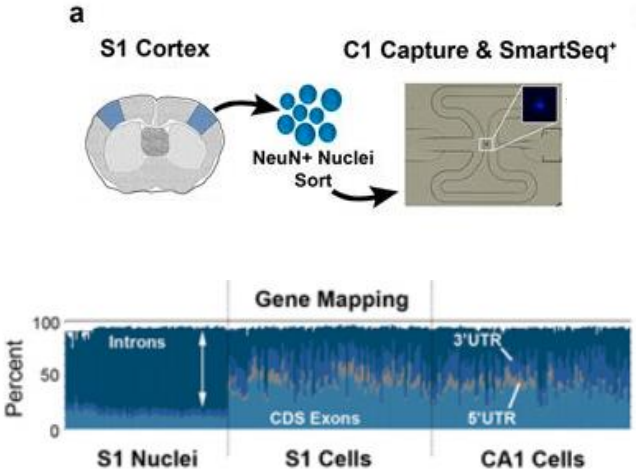
- Works for **frozen/post-mortem** data
- Nucleus vs cytosol
  - Lower genes count
  - Missing expression level of certain TFs
  - Intronic regions



Nucleus



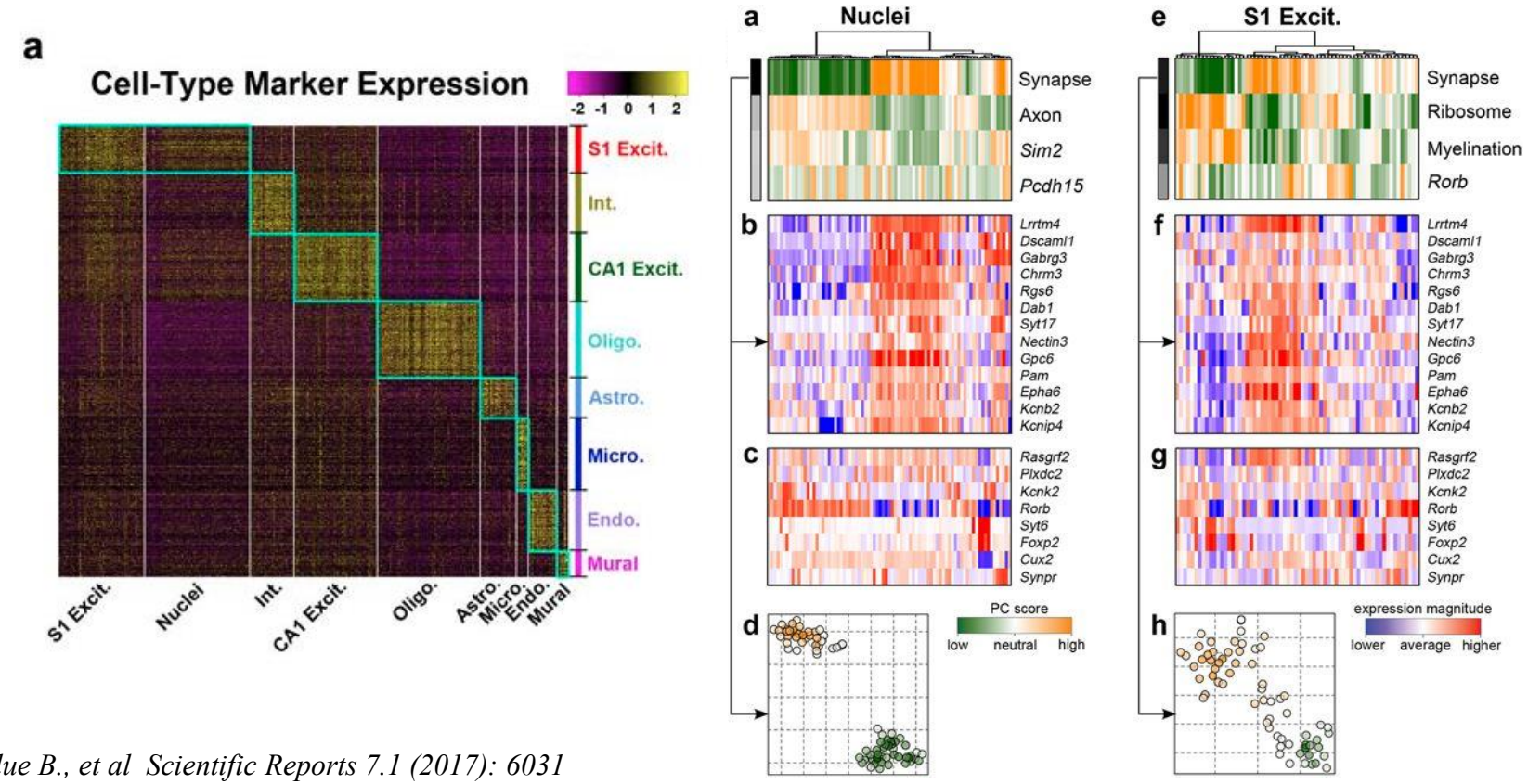




Lake, Blue B., et al Scientific Reports 7.1 (2017): 6031

Single nucleus RNA-seq

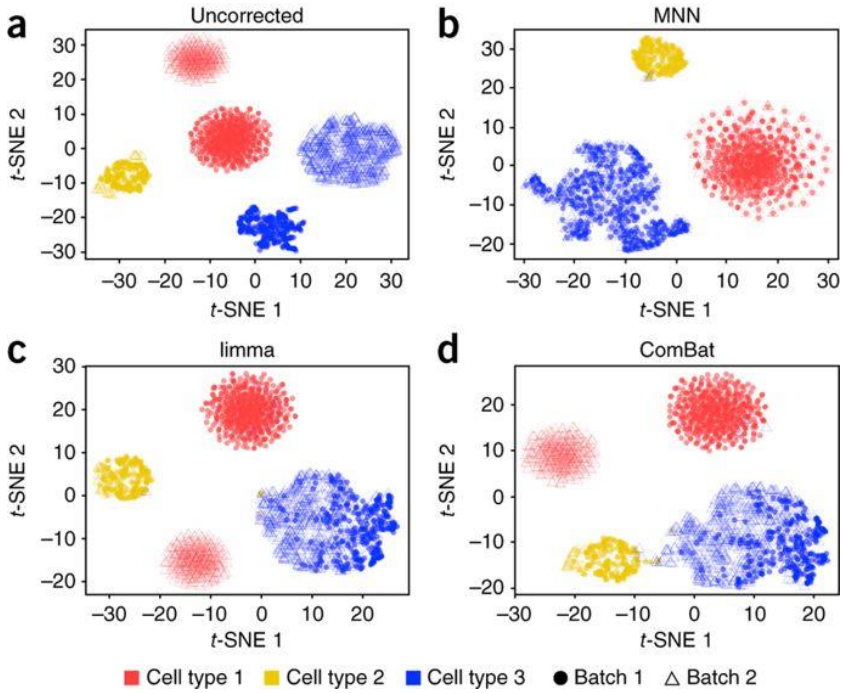
- **Biological similarity cell-nucleus** is confirmed …
- .. But different gene types also observed
- Requires careful preparation (different approaches)

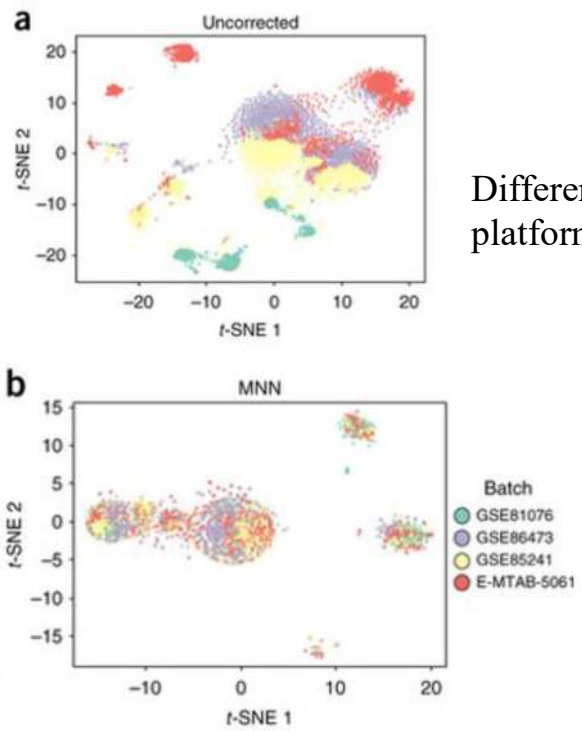


Lake, Blue B., et al Scientific Reports 7.1 (2017): 6031

Batch effect adjustment

- The distribution control does not bulk (ComBat or limma)
- Novel method: “mutual nearest neighbour”
  - Population shared between batches as control
  - Batch effect is smaller than biological effect



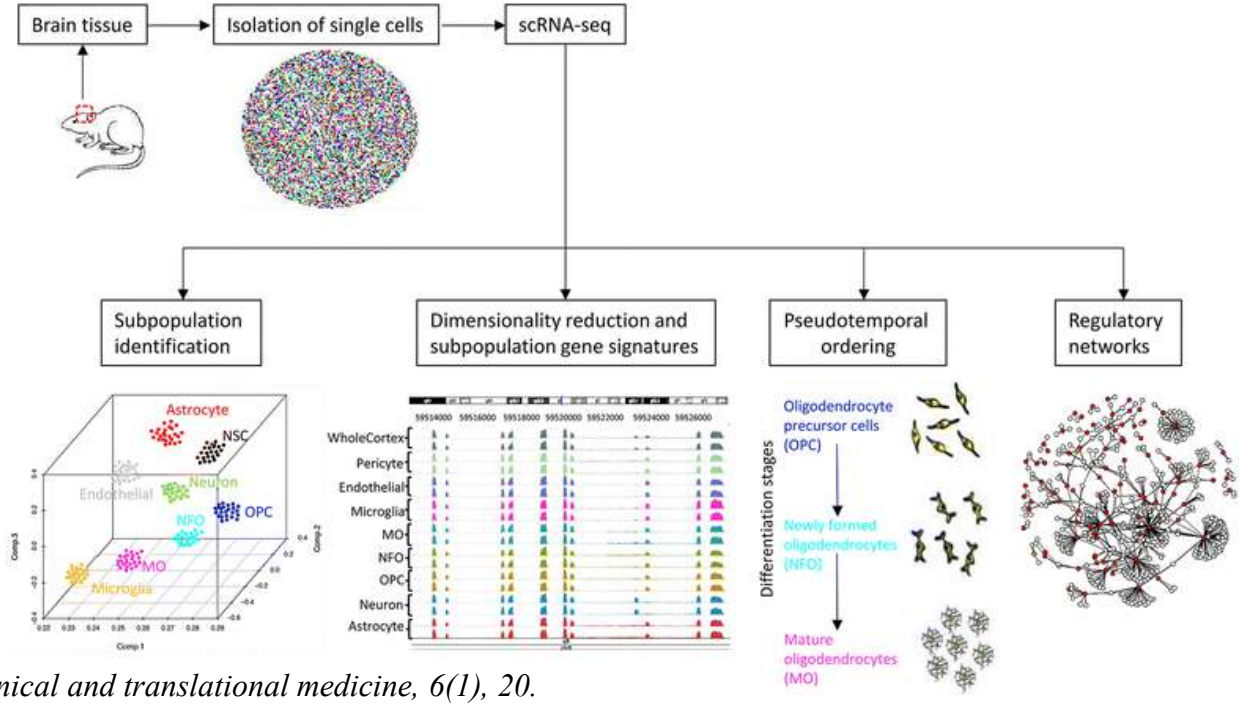


Different platforms

L. Haghverdi et al, Nature Biotechnology volume 36, pages 421–427 (2018)

Bioinformatics directions

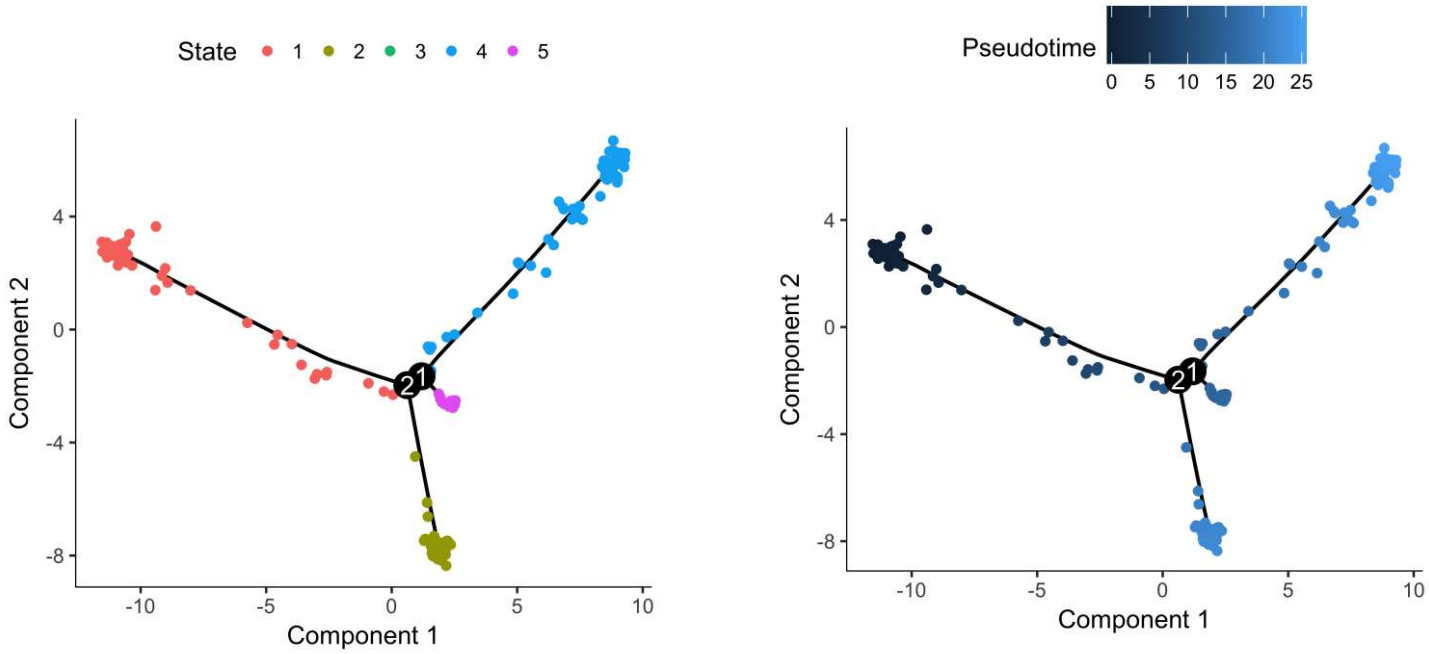
- Multiple **competitive areas** e.g. cell clustering to reduce dimensionality (PCA, tSNE, ZIFA,…)
  - Up to 20 methods already!
- Novel ideas are beneficial
- Biological confirmations are required



Duran, R. et al (2017). Clinical and translational medicine, 6(1), 20.

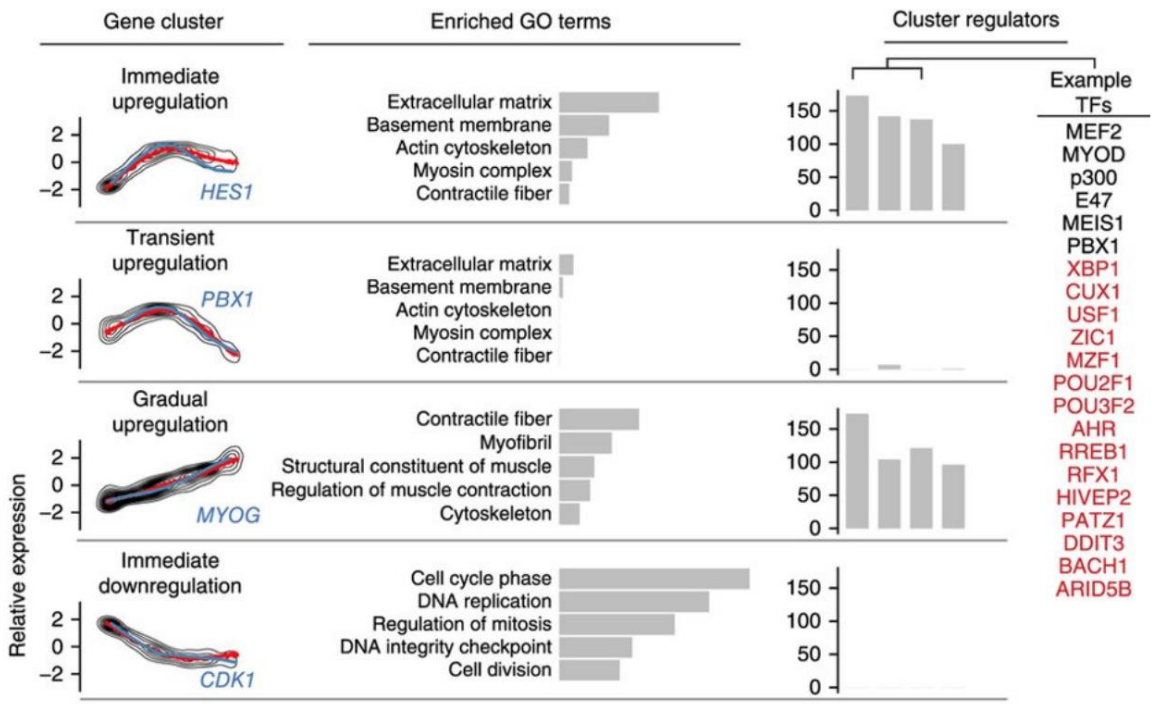
Pseudo-time and cell fate trajectories

- Expression patterns are unstable in cells: no strict type or no synchrony
- Transcriptional **reconfiguration**
  - Choose genes influencing the cell progress
  - Reduce dimensions, order cells along the trajectory



Pseudo-time and cell fate trajectories

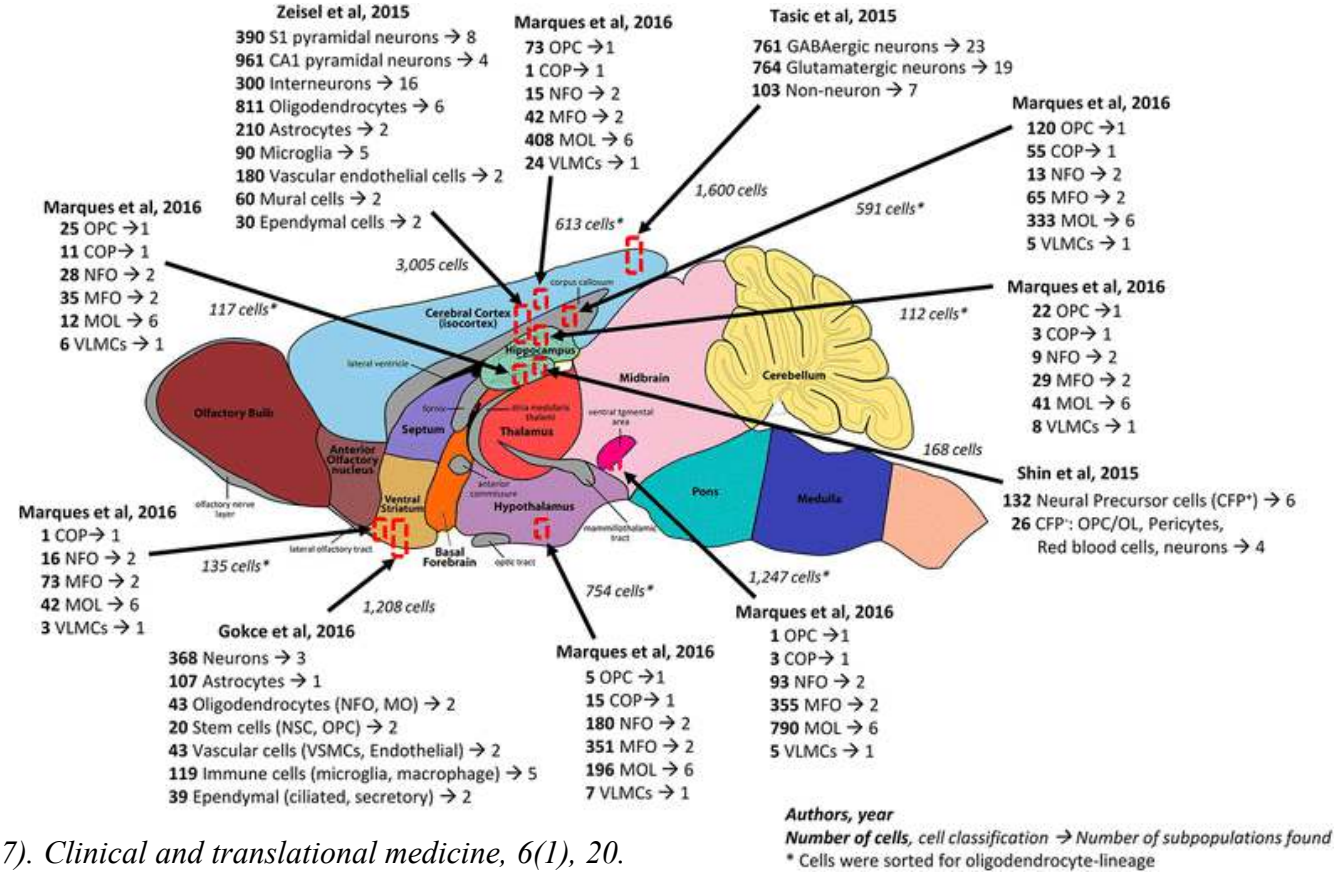
- Example study: dynamics of regulators for **cell fate** decision
- Activity of transcription factors regulating **differentiation**
  - Understand the phases of cells
  - Not possible through the bulk



C. Trapnell et al Nature Biotechnology 2014

Single cell RNA-seq: understand the brain

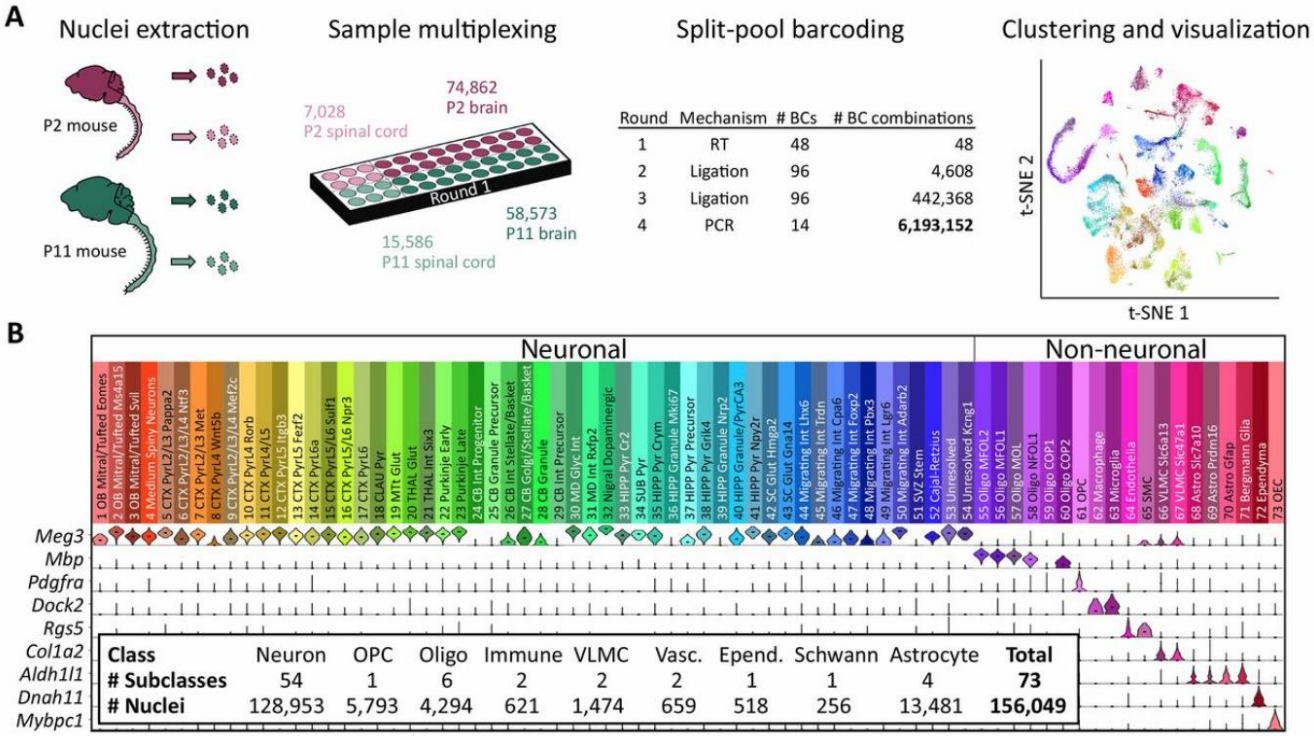
- Complex network of millions of diverse cells
- scRNA-seq: novel cell types, marker genes, networks
- Main focus - ***Mus Musculus,*** dream – ***Homo Sapiens***



Duran, R. et al (2017). Clinical and translational medicine, 6(1), 20.

Single cell RNA-seq: understand the brain

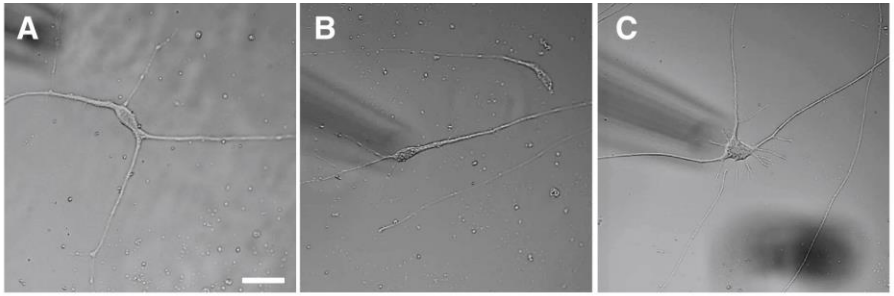
- Mouse full brain: ***split-pool ligation based*** transriptome sequencing
- Post-natal day 2 and 11, 156049 nucleus,100 cell types
- Central nervous system development



Rosenberg et al Science 15 Mar 2018

Single cell RNA-seq: understand the brain

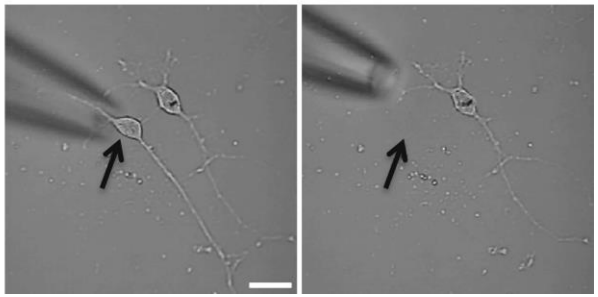
- Adult human brain: collect after neuro-surgery using ***microcapillary pipette*** (300 cells)
- Patient-specific transcriptional hierarchy:
  - Potential treatment detection
- Normal sequencing, pathway based analysis

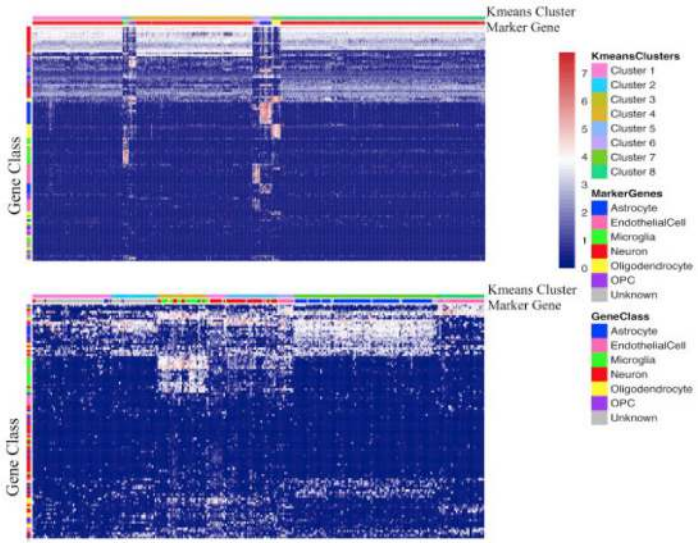


D

Pt.	A	G	E	FMAID	Diagnosis
22	24	M	C	242635	Epilepsy
50	63	F	C	270258	GBM
52	39	F	C	270342	Epilepsy
64	56	M	AA	270781	GBM
76	47	M	C	270797	GBM
82	32	F	C	270282	GBM
85	45	F	C	270348 275024	Epilepsy

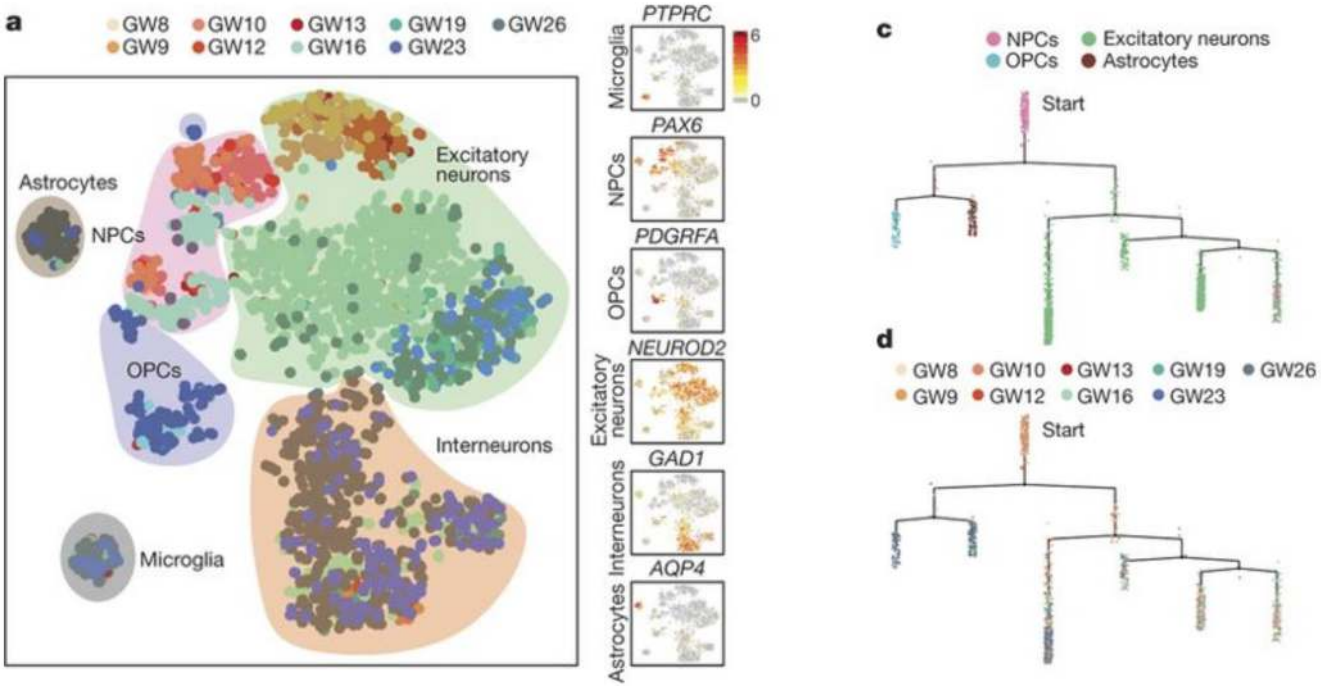
E





Single cell RNA-seq: understand the brain

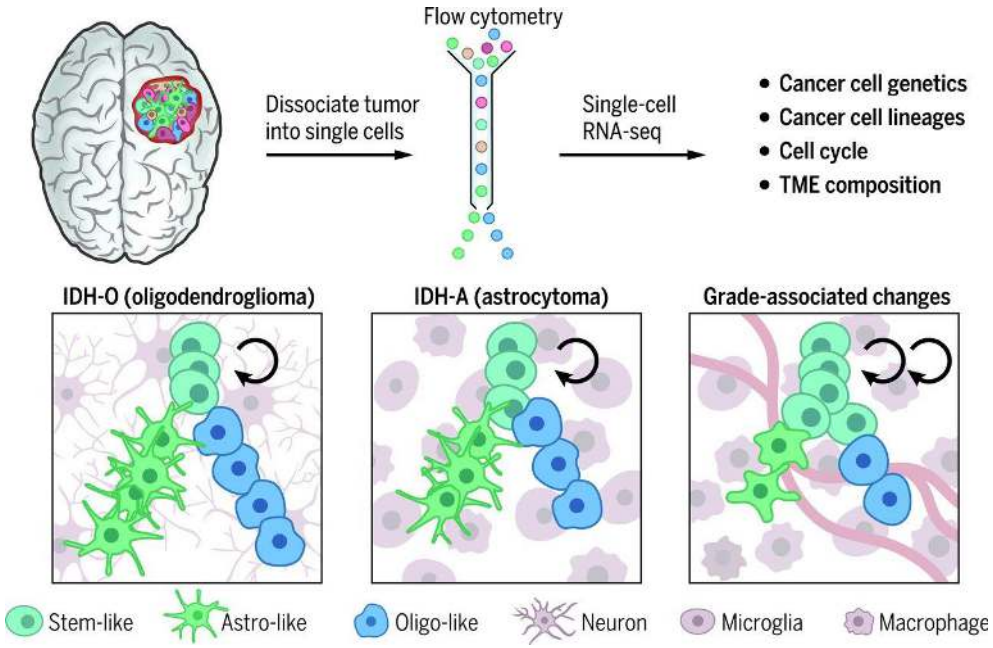
- Human prefrontal cortex: from 8 to 24 gestational weeks (study by Beijing Academy of Science)
- ***Smart-seq2*** protocol, 2300 cells, 35 cell types
- Dynamics of neurogenesis in the developing PFC



S. Zhong et al Nature volume 555, pages 524–528 (22 March 2018)

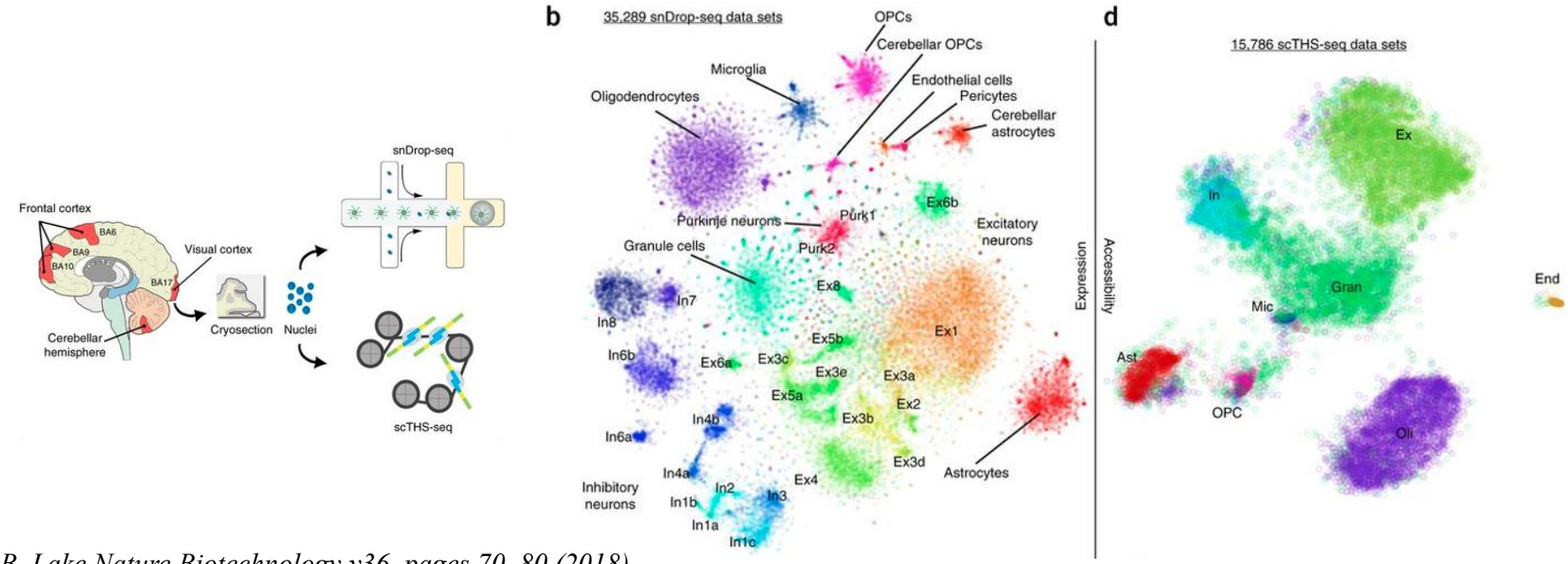
Single cell RNA-seq: brain tumors

- Brain cancer research:
  - Tumor subtypes cell lineage, development and growth
  - Cell of origin prediction
- Example: glioma. Different tumor types, but composoed in same cell types (7 patients, **Smart-seq C1**, ~4K cells)



Single cell multi-omics

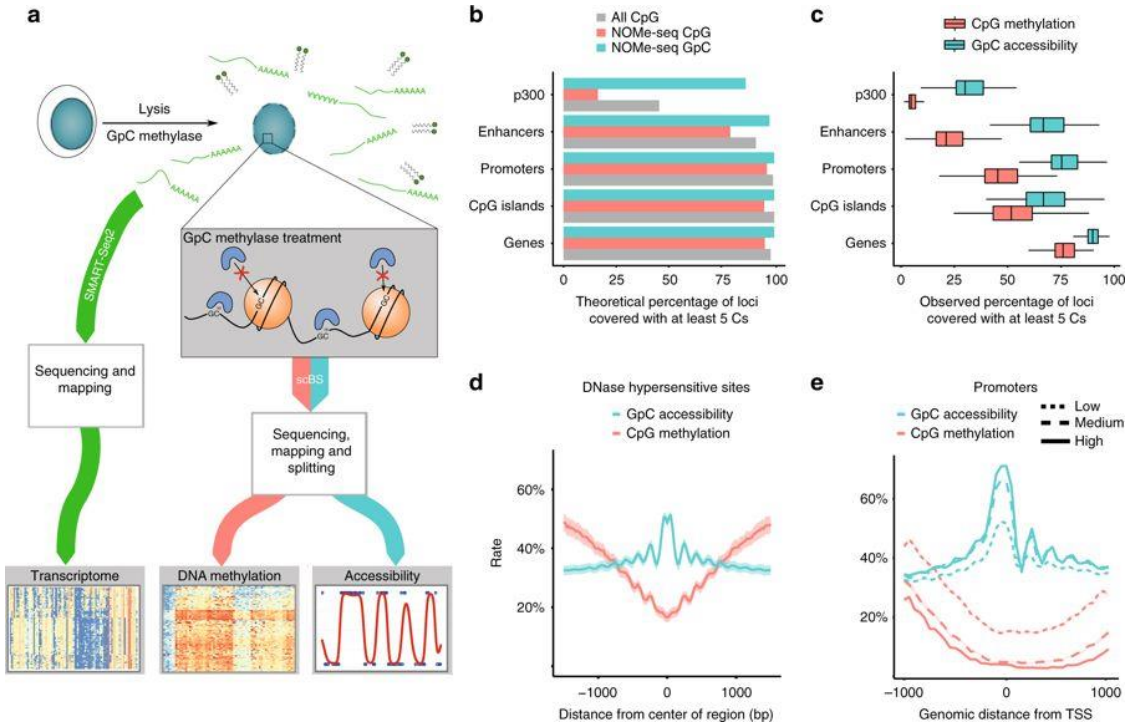
- Combination of data as in bulk studies
- Example: human visual and frontal cortex
  - snDROP-seq (frozen) : 16262 + 9794 cells
  - scTHS-seq (similar to ATAC-seq) : 13232 + 4753 cells
- Similar cell clusters structure was confirmed



B. Lake Nature Biotechnology v36, pages 70–80 (2018)

Single cell multi-omics

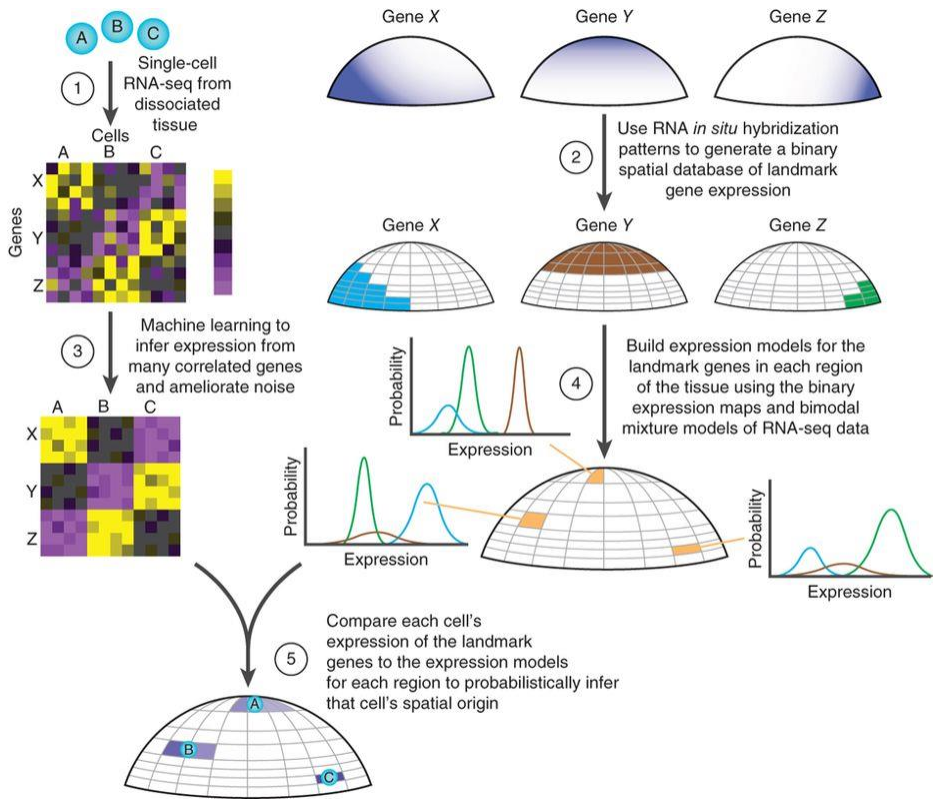
- More complex combination: same cells usage
  - Transcriptome, methylation, chromatin accessibility
- Allows to undesrtand the individual genomic loci in a cell
  - Hypomethylated : open, hyper : closed

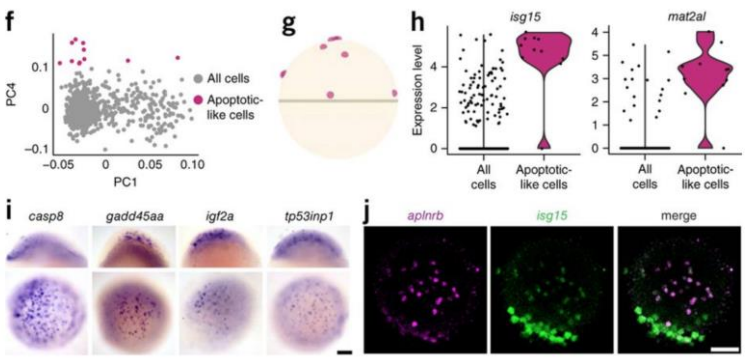


S. J. Clark Nature Communications v9, number: 781 (2018)

Single cell spatial control

- Spatial localization important for celluar fate
- Solution: integrate snRNA-seq and in-situ data (e.g. FISH)
  - Expression patterns reveal spatial archetypes



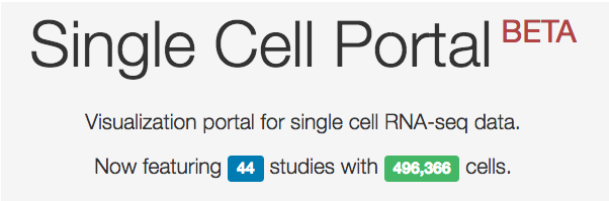


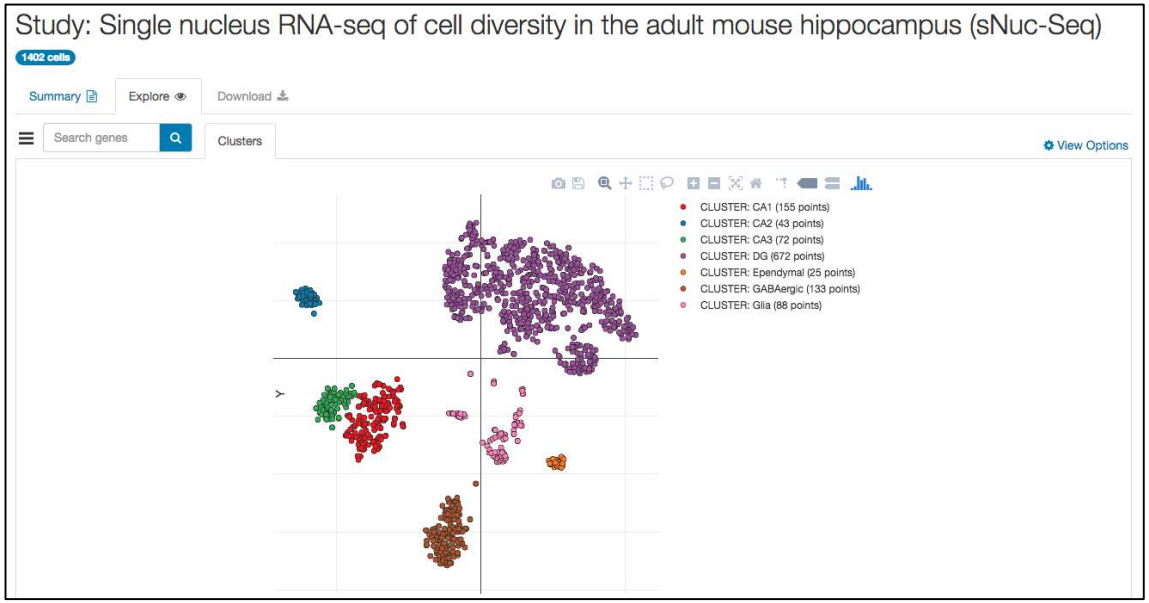
- *Correct spatial organization*
- *Characterization of rare cell populations*

R. Satija Nature Biotechnology volume 33, pages 495–502 (2015)

BRAIN initiative Cell Census Consortium

- Detailed collection of existing data
  - Brain anatomical atlas creation
  - Species comparison
- Validation methods estimation e.g. microscopy and labelling
- Single cell portal: *portals.broadinstitute.org/single_cell*
  - Data availability





Human Cell Atlas

- Comprehensive reference maps of **all human cells**
- Human health understanding for diagonsing, monitoring and treatment
- Areas of impact: immune system, brain, epithelial tissue, cancer, developmental cells
- Notable people: Aviv Regev, Sarah Teichmann, Sten Linnarson, John Marioni …
- A lot of politics

HUMAN
CELL
ATLAS

Other sources

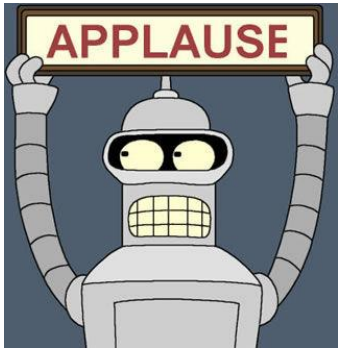
- Notable example: **Chan Zuckerberg** initiative
  - Human Cell Atlas financial support

CHAN ZUCKERBERG INITIATIVE

- **Single Cell Genomics Day 2018**
  - Live workshop, fully recorded recorded. Topics covered:
    - Single nucleus sequencing
    - Batch effect correction
    - Multi-modal analysis
    - Multiplexing single cell experiments
    - RNA velocity
    - Split-pool combinatorial barcoding

Summary

- Single cell sequencing (genome, RNA, etc) is an **effective** method to understand the cells activity, but **has certain limits**
- Problematic issues: **low proportion** of the genes present, **limited coverage** of transcript and dealing with **frozen** tissue. All of them can be handled by careful adjustment.
- Multiple benefits of usage can be obtained from **novel computational methods** such as for example **pseudotime** control
- Frequently updated **public resources** such as for example Single Cell Portal are already available to dig in.

Спасибо за внимание!

Вопросы?