BROTMAN BATY INSTITUTE

Single-cell RNA-Sequencing Introduction

Mary O'Neill, Ph.D.
Director, Single Cell Genomics
ONeillMB@uw.edu
@ONeillMB1

Anh Vo, M.S.
Bioinformatician
athuyvo@uw.edu

BBI Advanced Technology Lab – Single Cell Genomics Group

Kathryn Barker
RESEARCH SCIENTIST 1

Brent Ewing
SOFTWARE ENGINEER

Diana O'Day
RESEARCH SCIENTIST 3

Mary O'Neill
DIRECTOR, SINGLE CELL GENOMICS

Anh Vo
GENOMIC DATA ANALYST

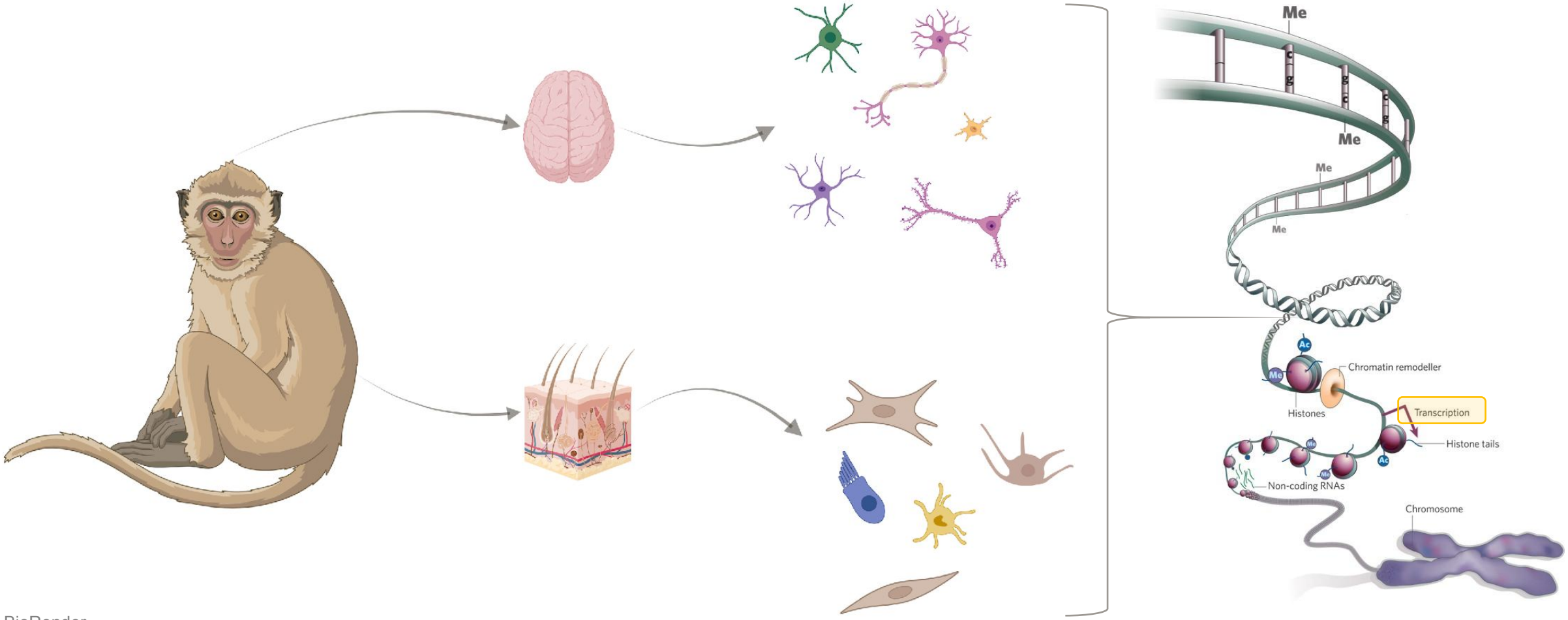
Luis Gamboa
RESEARCH SCIENTIST 2

Disclosures

- BBI has a fee-for-service single-cell sequencing platform
    - We offer non-commercial combinatorial indexing methods
    - We work closely with the Shendure Lab who developed sci-RNA-seq and the Trapnell Lab who develop Monocle3
- In the past, some of us have worked with 10X Genomics 3’ Gene Expression
    - Performed most analyses in the Bioconductor Single Cell Universe (e.g. Scater, Scran, Scuttle, DropletUtils)

BROTMAN BATY
INSTITUTE

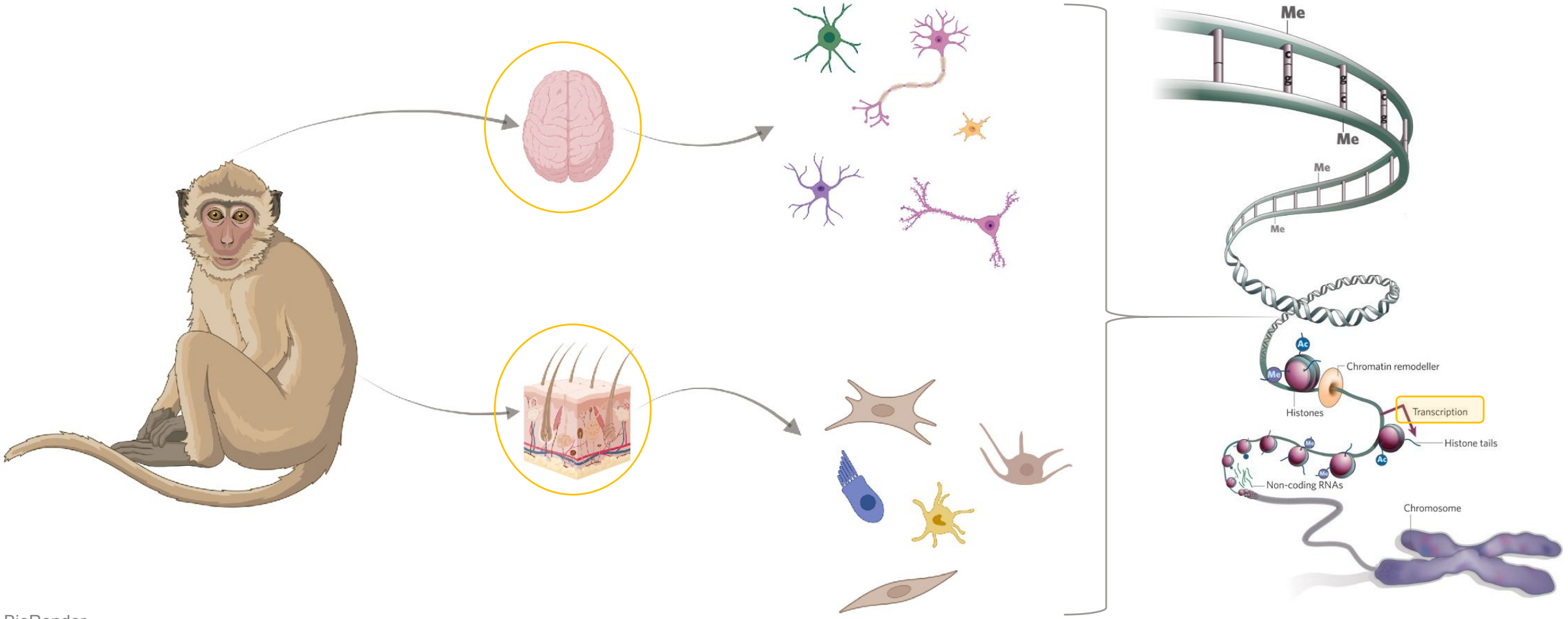
Incredible cellular diversity encoded in the genome



BioRender

https://doi.org/10.1038/454711a

Bulk methods average profiles across all cells of a sample

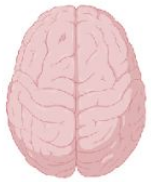


BioRender

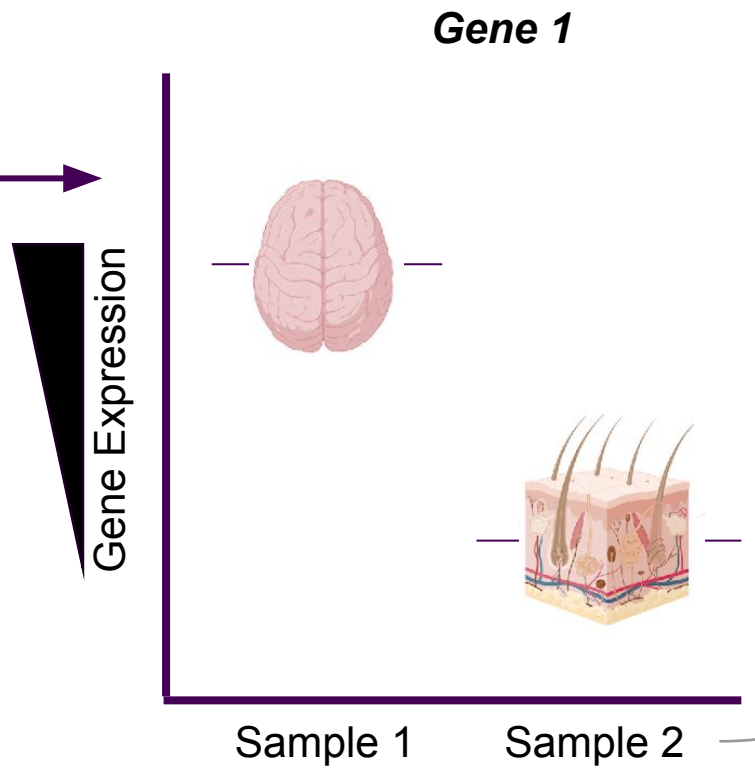
https://doi.org/10.1038/454711a

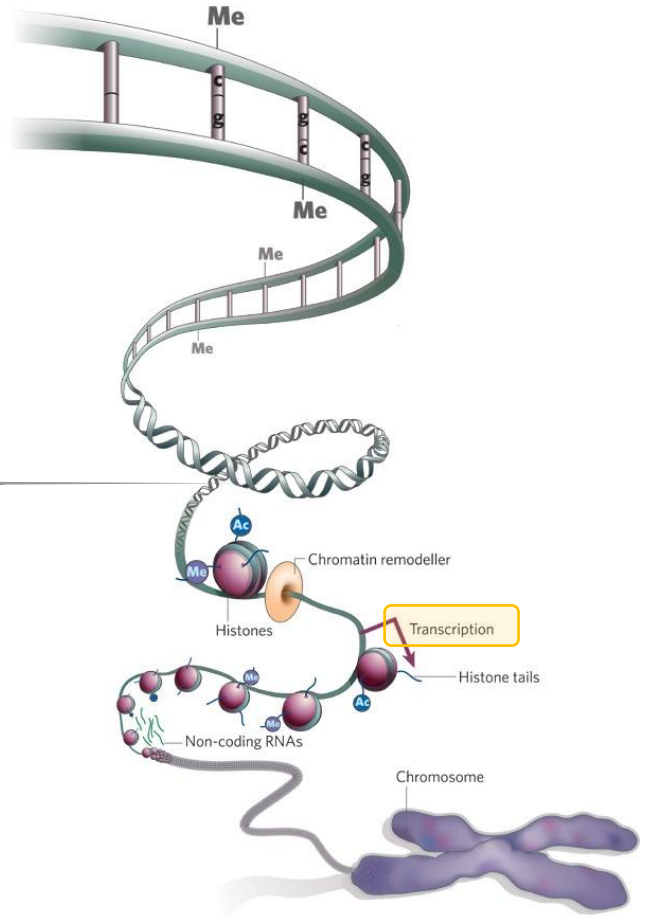
Bulk methods average profiles across all cells of a sample

e.g. RNA-seq



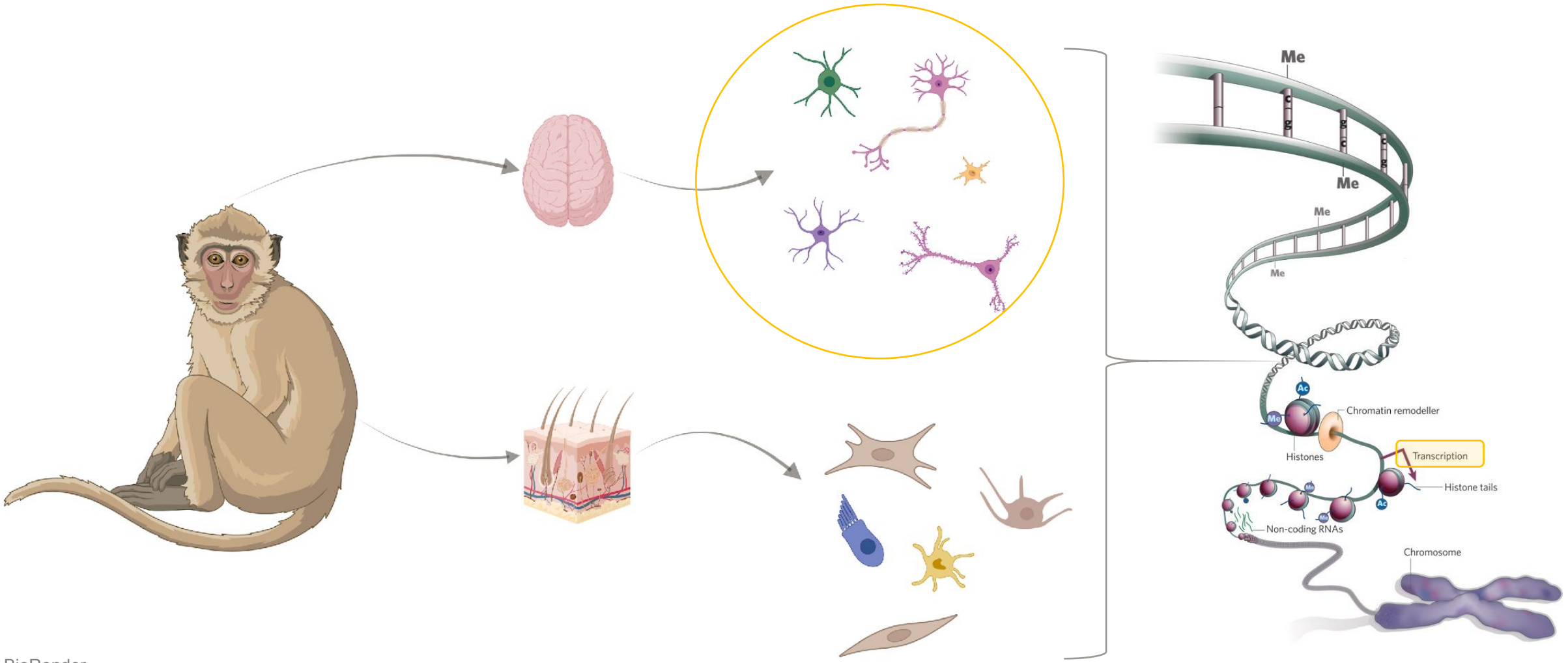
	Sample 1	Sample 2
Gene 1	2000	100
Gene 2	1000	1010
Gene 3	3500	300
Gene 4	50	4000
Gene 5	3	5
...	...	...
Gene n	500	2400





https://doi.org/10.1038/454711a

But what if you are actually interested in this?



BioRender

https://doi.org/10.1038/454711a

Single cell methods provide cellular resolution

e.g. scRNA-seq

This can be hundreds to millions of cells

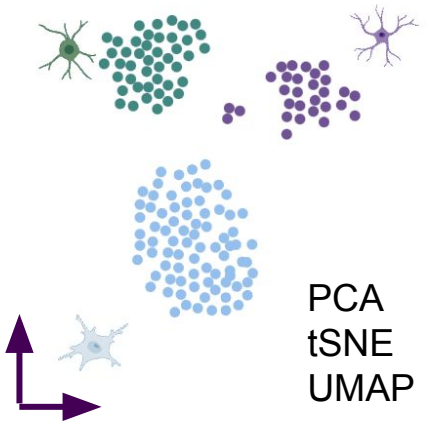
	Cell 1	**Cell 2**	**Cell 3**	**…**	**Cell n**
Gene 1	30	2	21	…	20
Gene 2	11	10	14	…	14
Gene 3	12	35	6	…	5
Gene 4	0	0	0	…	0
Gene 5	0	0	0	…	0
…	…	…	…	…	…
Gene n	1	10	2	…	2

A lot of zeros

Single cell methods provide cellular resolution

e.g.
scRNA-seq

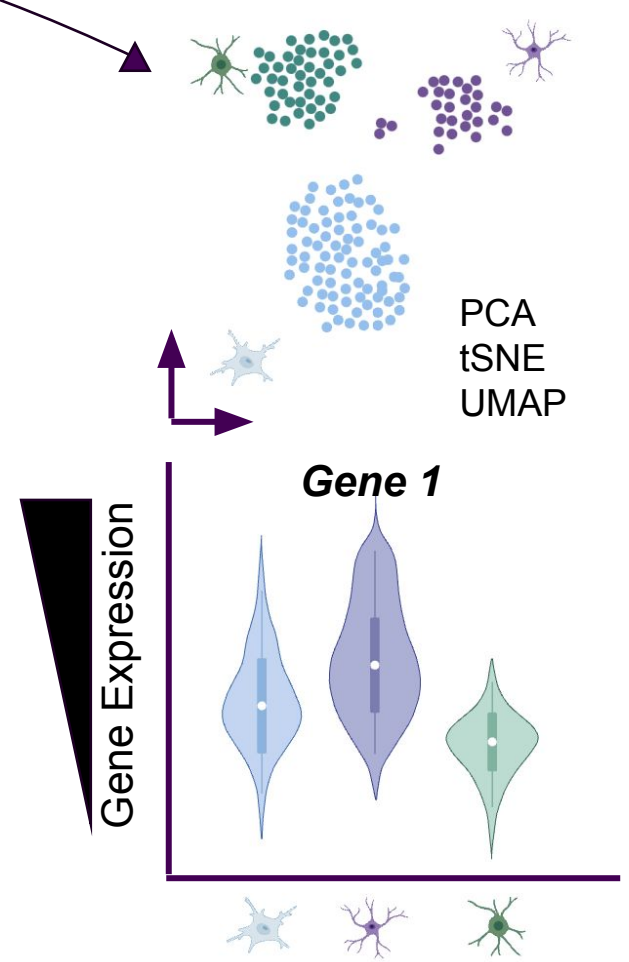
	Cell 1	Cell 2	Cell 3	…	Cell n
Gene 1	30	2	21	…	20
Gene 2	11	10	14	…	14
Gene 3	12	35	6	…	5
Gene 4	0	0	0	…	0
Gene 5	0	0	0	…	0
…	…	…	…	…	…
Gene n	1	10	2	…	2



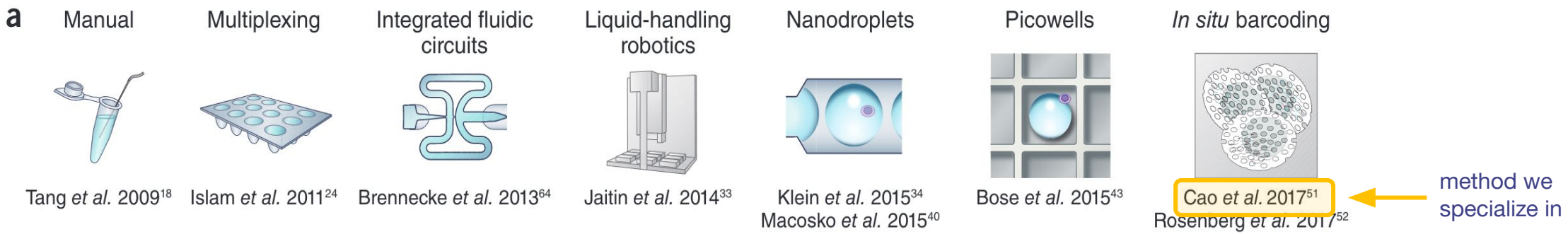
Single cell methods provide cellular resolution

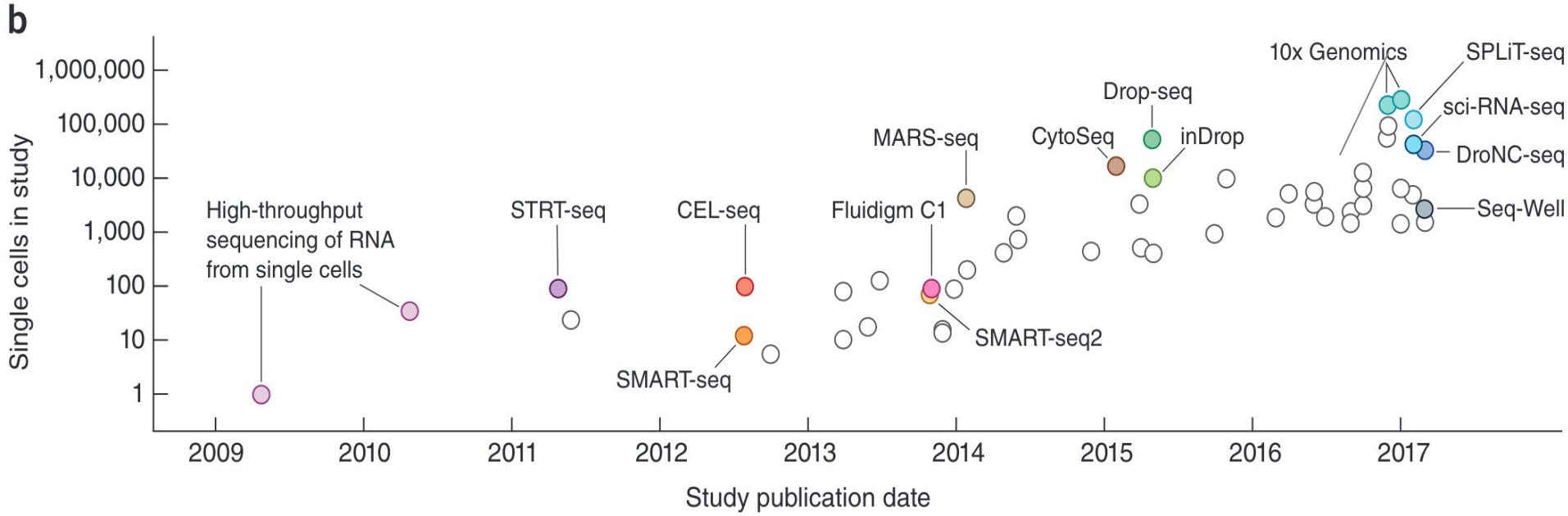
e.g. scRNA-seq

	Cell 1	Cell 2	Cell 3	…	Cell n
Gene 1	30	2	21	…	20
Gene 2	11	10	14	…	14
Gene 3	12	35	6	…	5
Gene 4	0	0	0	…	0
Gene 5	0	0	0	…	0
…	…	…	…	…	…
Gene n	1	10	2	…	2



Timeline & throughput of scRNA-seq

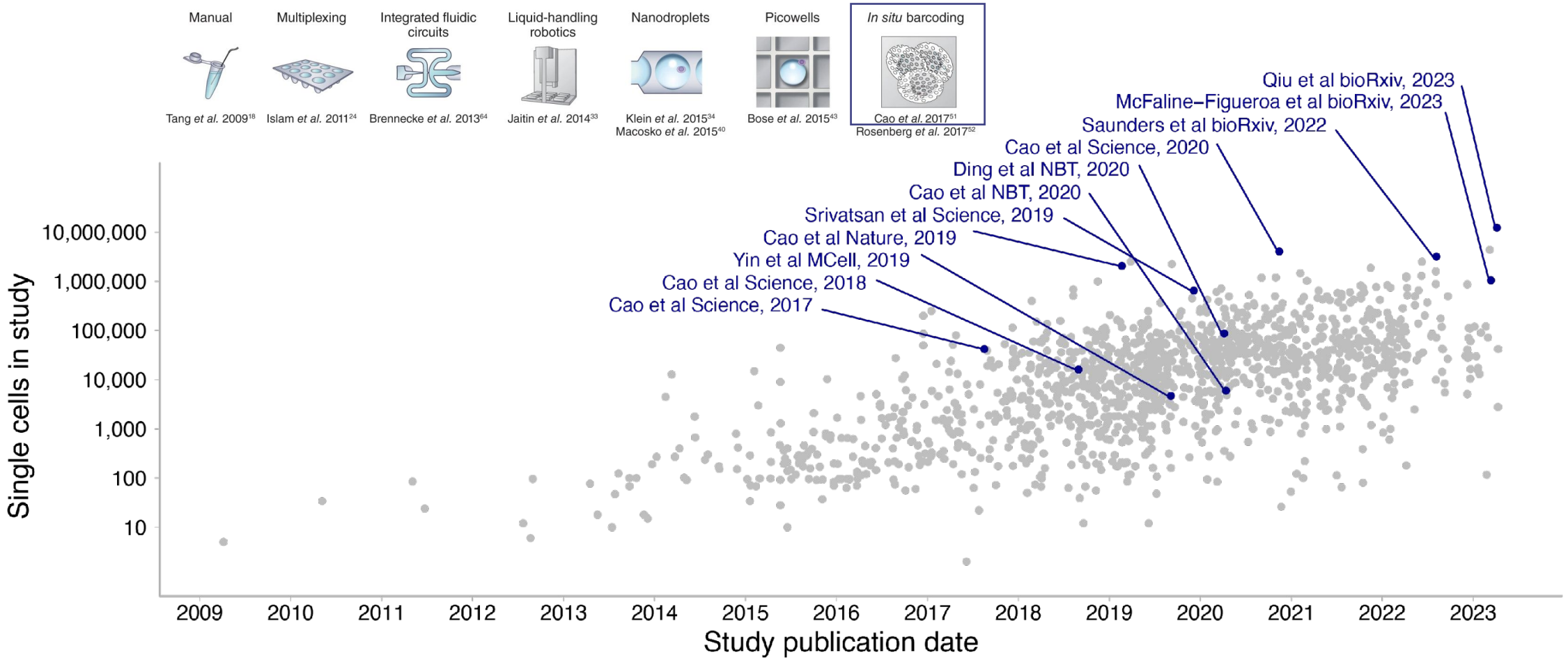




Svensson et al. 2018 (Nature Protocols)

Scaling of sci-RNA-seq experiments

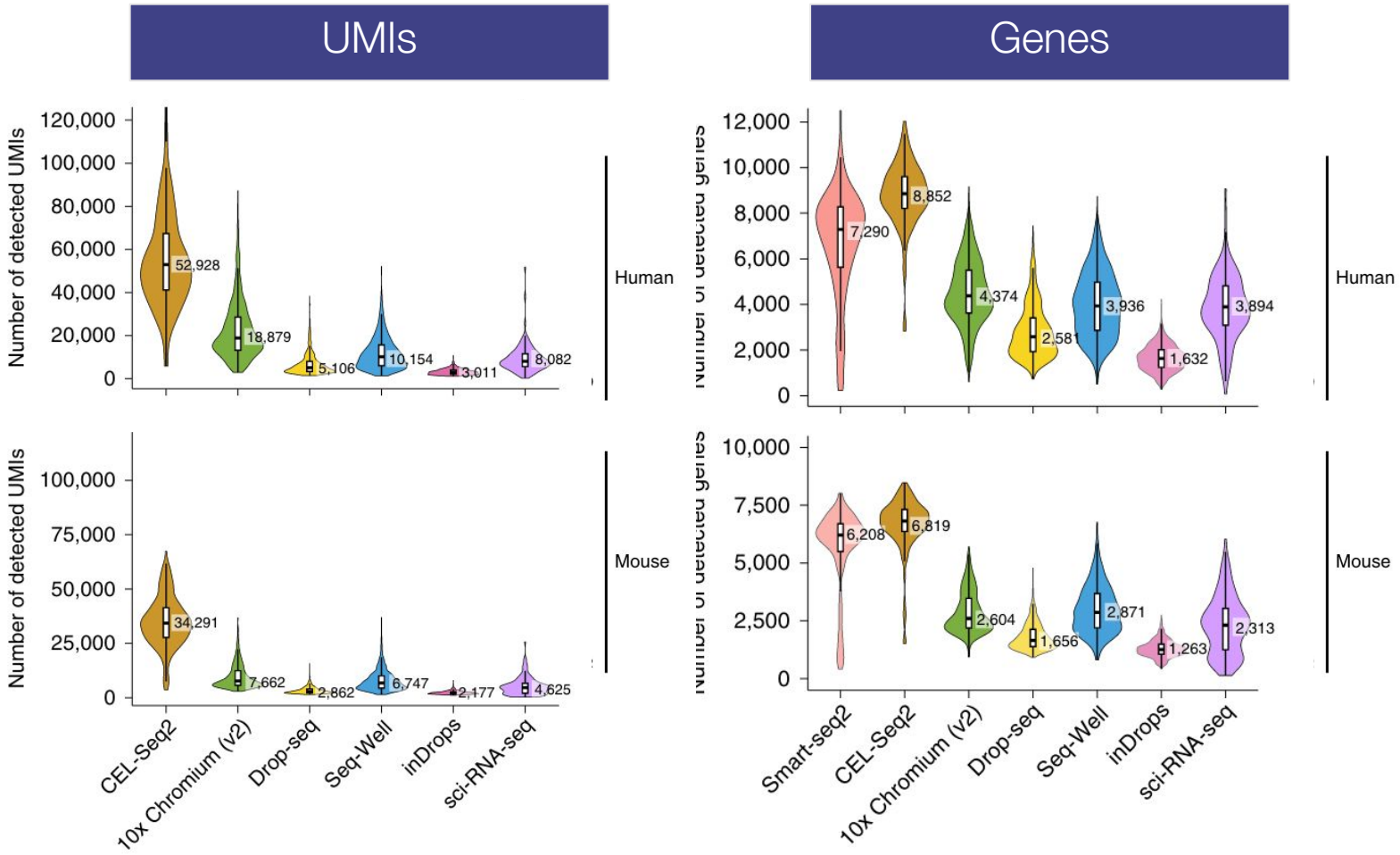
Single-cell combinatorial indexing (sci-)



Svensson *et al.* 2018 (Nature Protocols)
Data from 'Single cell studies database', Svensson 2023

RNA capture / sensitivity

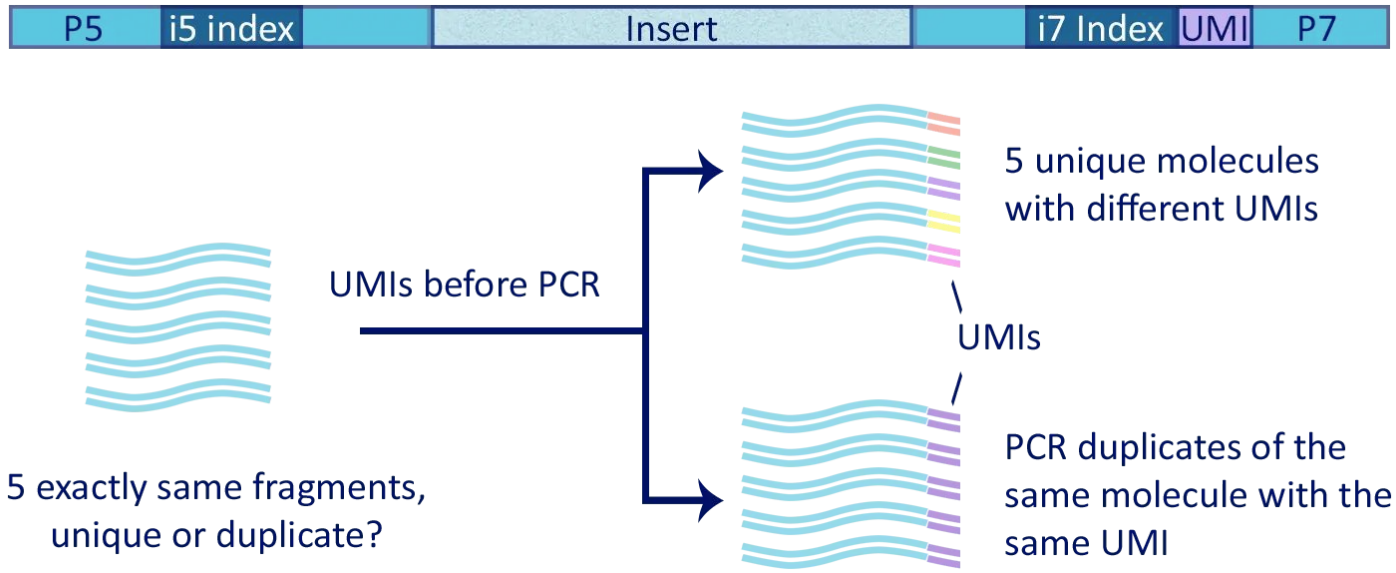
What is a UMI?



Ding et al. 2020 (Nature Biotechnology)

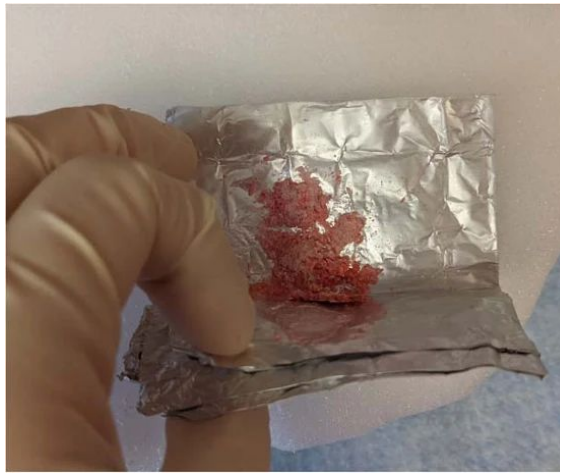
What is a UMI?

- Unique Molecular Identifier (UMI)
- Low input material may cause amplification bias
- UMIs are sequences that correspond to one fragment
- Sequenced reads with the same UMI are from the same fragment
- Unique sequenced collapsed to a single for counting



Single-cell vs single-nuclei

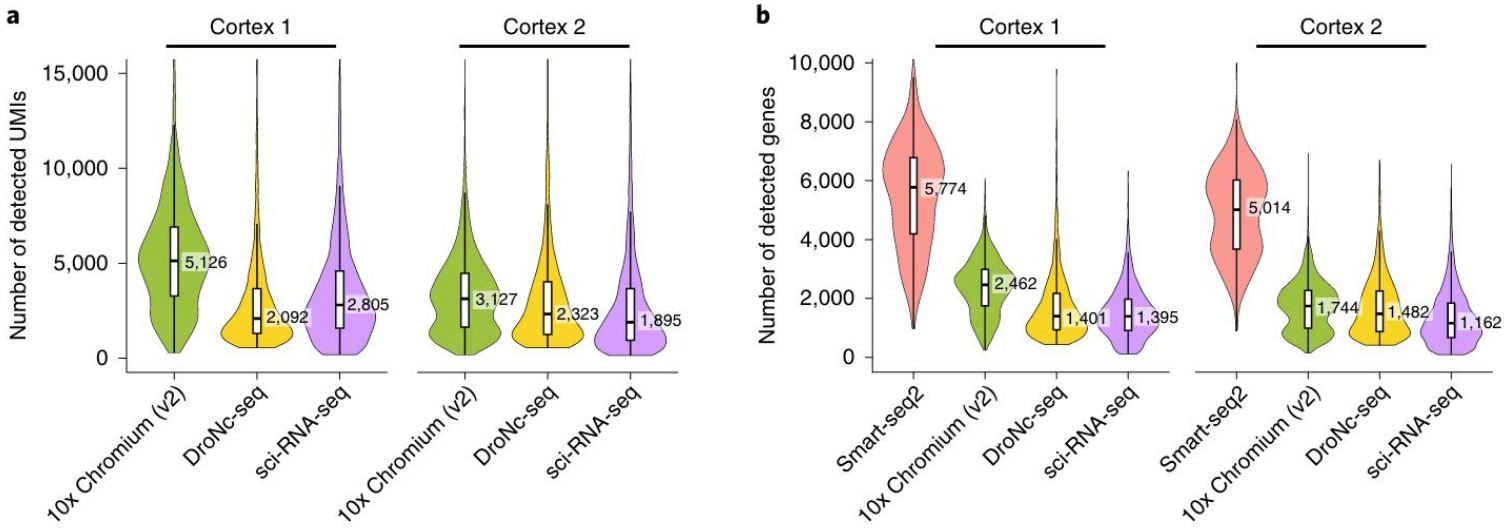
- Single-cell profiling does not always provide an unbiased view on cell types
- Some cell types are more vulnerable to the tissue dissociation process
- For example, brain. Glutamatergic neurons more sensitive, non-neuronal cells overrepresented
- Single-cell sequencing often relies on fresh tissue
- Nuclei more resistant to mechanical force, safely isolated from frozen tissue w/o dissociation enzymes



RNA capture / sensitivity

UMIs

Genes



Ding et al. 2020 (Nature Biotechnology)

Other considerations in choosing a technology

	SMART-seq2	CEL-seq2	STRT-seq	Quartz-seq2	MARS-seq	Drop-seq	inDrop	Chromium	Seq-Well	sci-RNA-seq	SPLiT-seq
Single-cell isolation	FACS, microfluidics	FACS, microfluidics	FACS, microfluidics, nanowells	FACS	FACS	Droplet	Droplet	Droplet	Nanowells	Not needed	Not needed
Second strand synthesis	TSO	RNase H and DNA pol I	TSO	PolyA tailing and primer ligation	RNase H and DNA pol I	TSO	RNase H and DNA pol I	TSO	TSO	RNase H and DNA pol I	TSO
Full-length cDNA synthesis?	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes
Barcode addition	Library PCR with barcoded primers	Barcoded RT primers	Barcoded TSOs	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers and library PCR with barcoded primers	Ligation of barcoded RT primers
Pooling before library?	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Library amplification	PCR	In vitro transcription	PCR	PCR	In vitro transcription	PCR	In vitro transcription	PCR	PCR	PCR	PCR
Gene coverage	Full-length	3′	5′	3′	3′	3′	3′	3′	3′	3′	3′
Number of cells per assay											

Chen X, et al. 2018.
Annu. Rev. Biomed. Data Sci. 1:29–51

Typical scRNA-seq protocol

- Tissue dissection and cell dissociating to obtain a suspension of cells.
- Optionally cells may be selected (e.g. based on membrane markers, fluorescent transgenes or staining dyes).
- Capture single cells into individual reaction containers (e.g. wells or oil droplets). Note, for combinatorial indexing methods the cell/nucleus itself serves as the reaction chamber.
- Extracting the RNA from each cell.
- Reverse-transcribing the RNA to more stable cDNA.
- Amplifying the cDNA (either by *in vitro* transcription or by PCR).
- Preparing the sequencing library with adequate molecular adapters.
- Sequencing, usually with paired-end Illumina protocols.
- Processing the raw data to obtain a count matrix of genes-by-cells
- Carrying-out several downstream analysis

10X Flavor

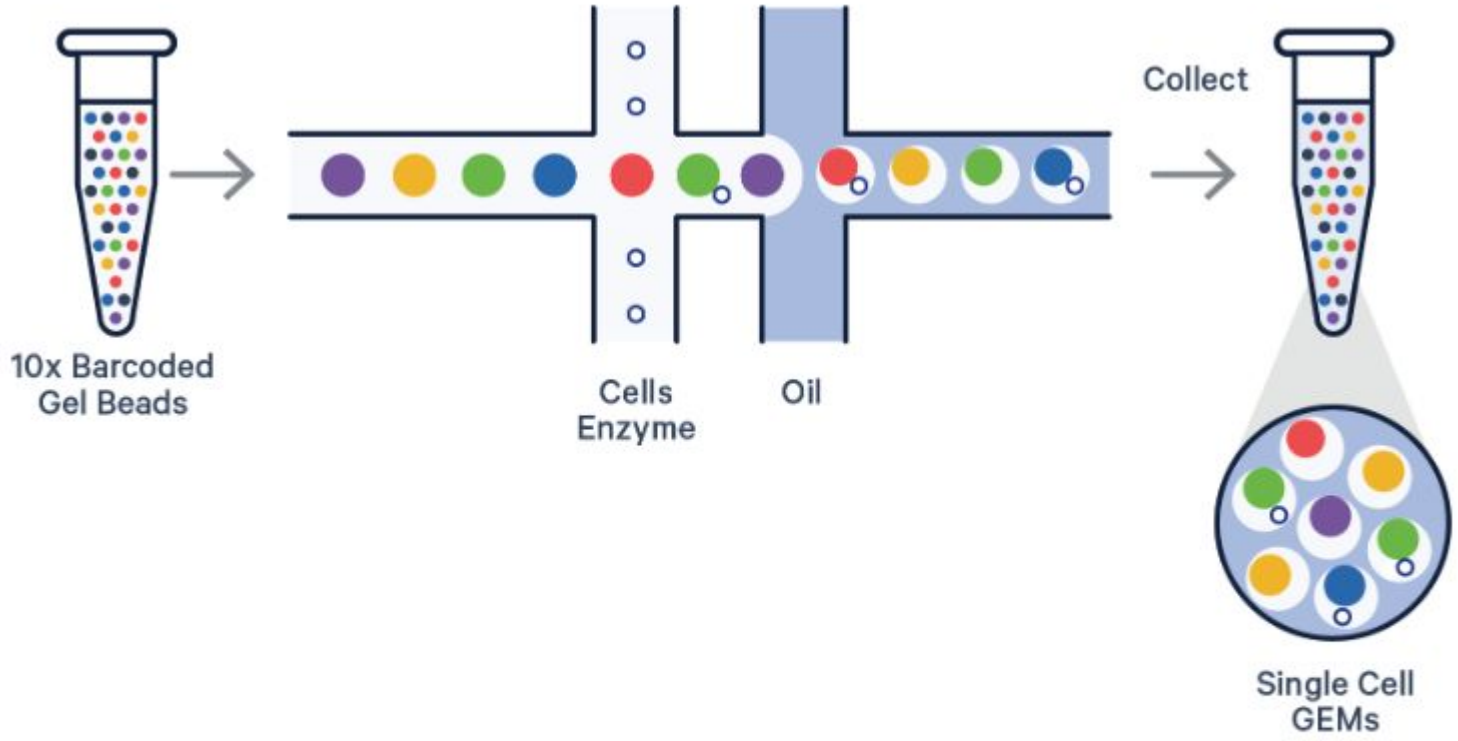


Figure 1. A schematic overview of GEM generation and barcoding with the GEM-X chip workflow. GEMs are generated by combining barcoded Gel Beads, a master mix containing cells, and partitioning oil in a GEM-X 3' or 5' Chip. To achieve single cell resolution, cells are delivered at a limiting dilution, such that the majority (~90–99%) of generated GEMs contain no cell, while the remainder largely contain a single cell.

10X Flavor

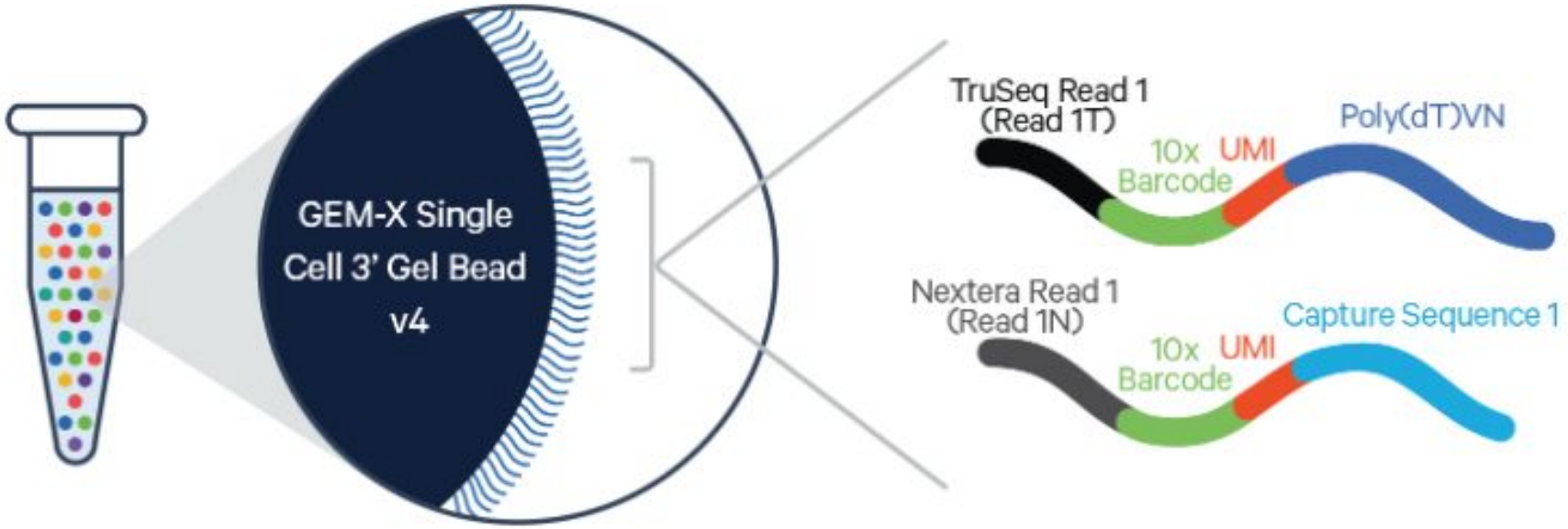
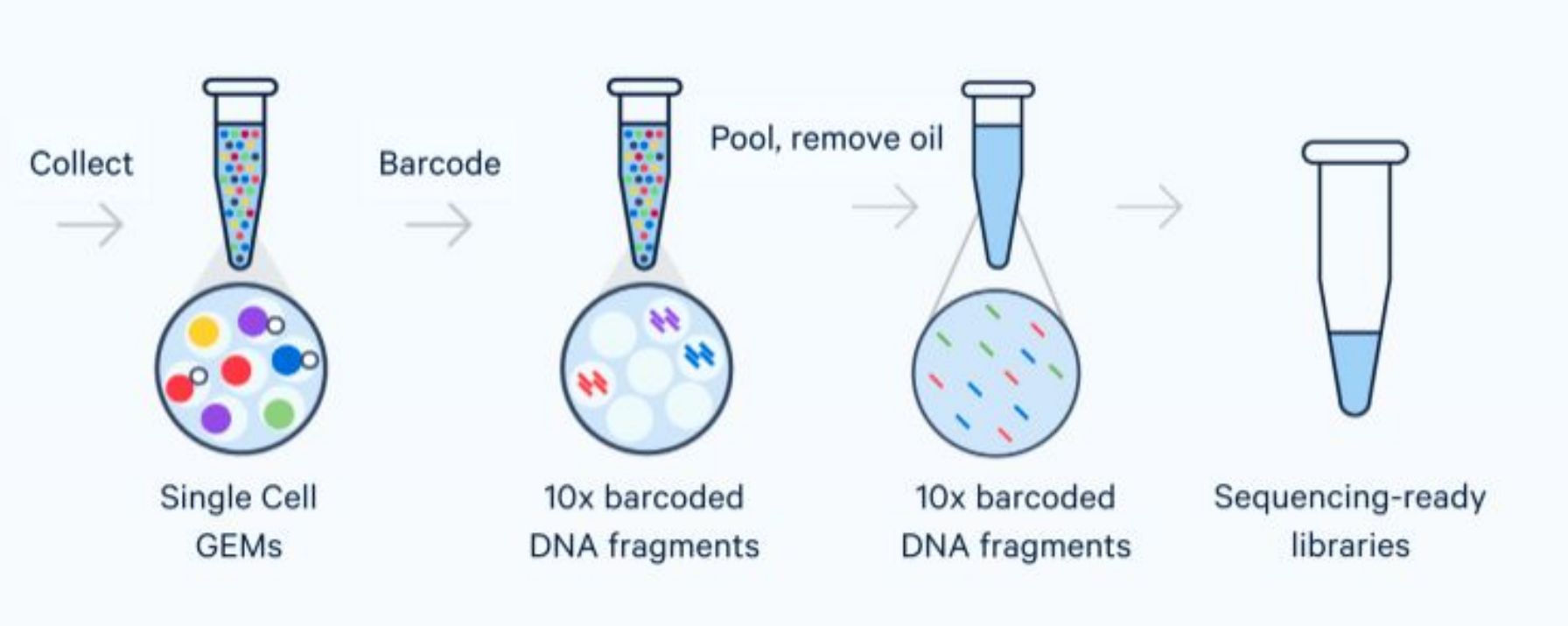
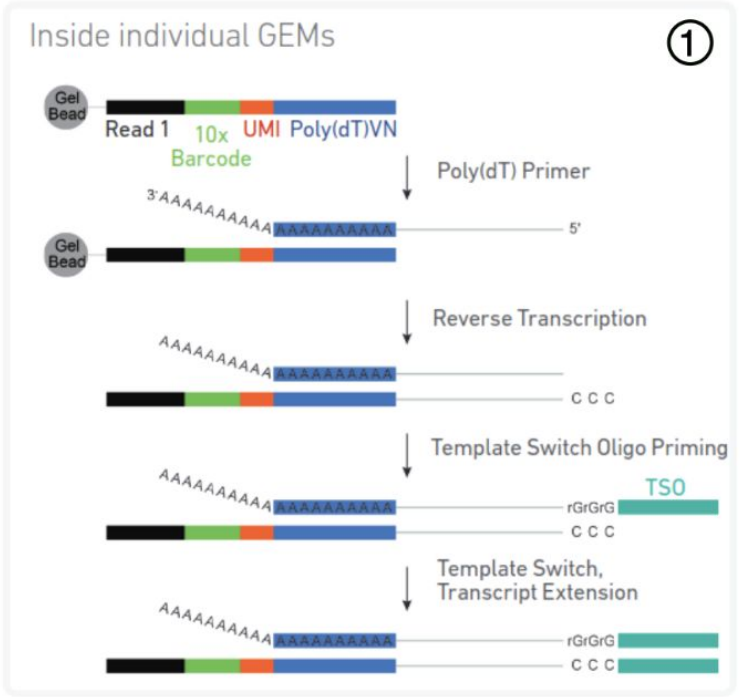


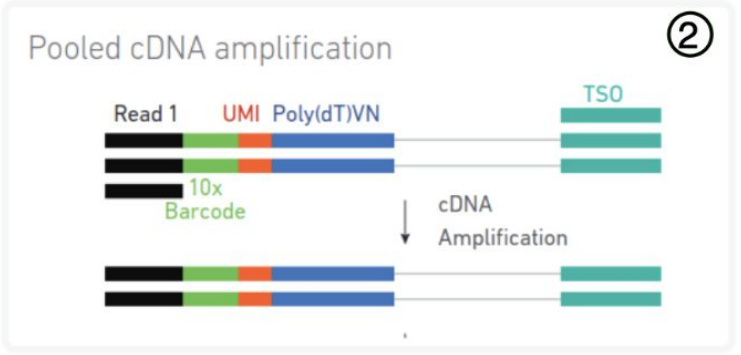
Figure 2. Schematic diagram of a GEM-X Single Cell 3' Gel Bead. Every Gel Bead is coated with oligos containing an Illumina TruSeq Read 1 (read 1 sequencing primer, Read 1T), 16 nt 10x Barcode, 12 nt unique molecular identifier (UMI), and 30 nt poly(dT)VN.

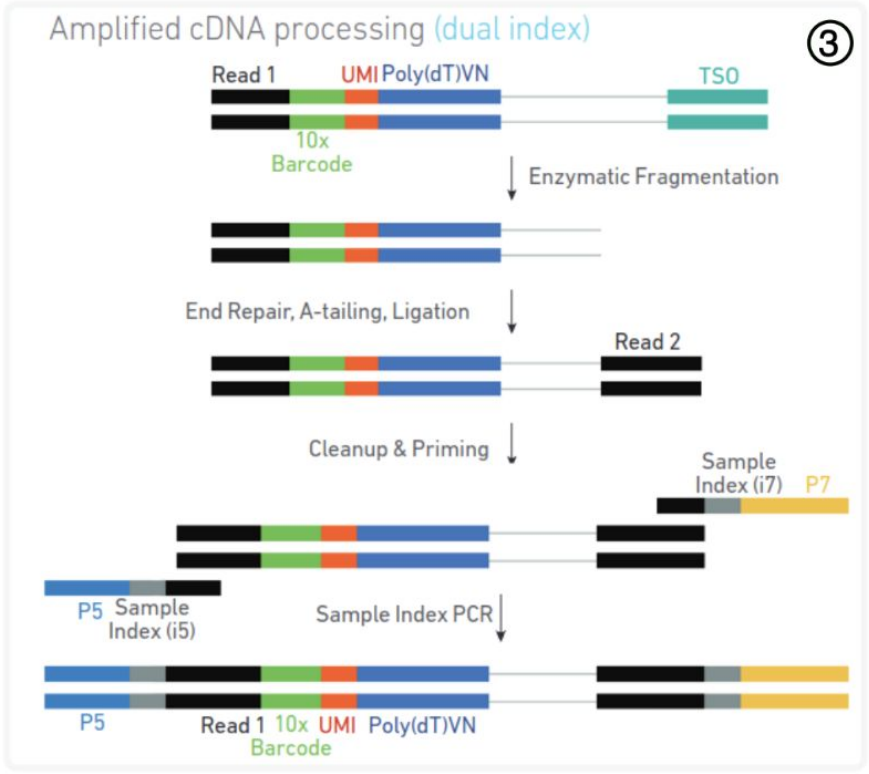
10X Flavor

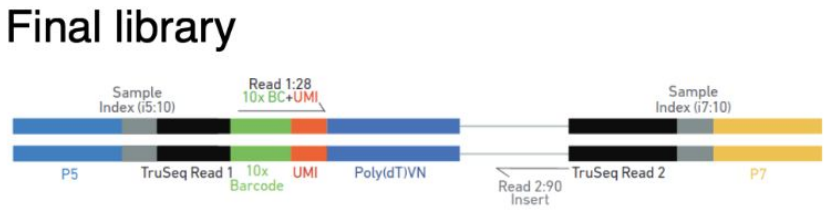

Collect
Barcode
Pool, remove oil
Single Cell GEMs
10x barcoded DNA fragments
10x barcoded DNA fragments
Sequencing-ready libraries

10X Flavor


Inside individual GEMs
①
Gel Bead
Read 1
10x Barcode
UMI
Poly(dT)VN
Poly(dT) Primer
3'
5'
Reverse Transcription
C C C
Template Switch Oligo Priming
TSO
rGrGrG
Template Switch, Transcript Extension


Pooled cDNA amplification
②
Read 1
UMI
Poly(dT)VN
TSO
10x Barcode
cDNA Amplification

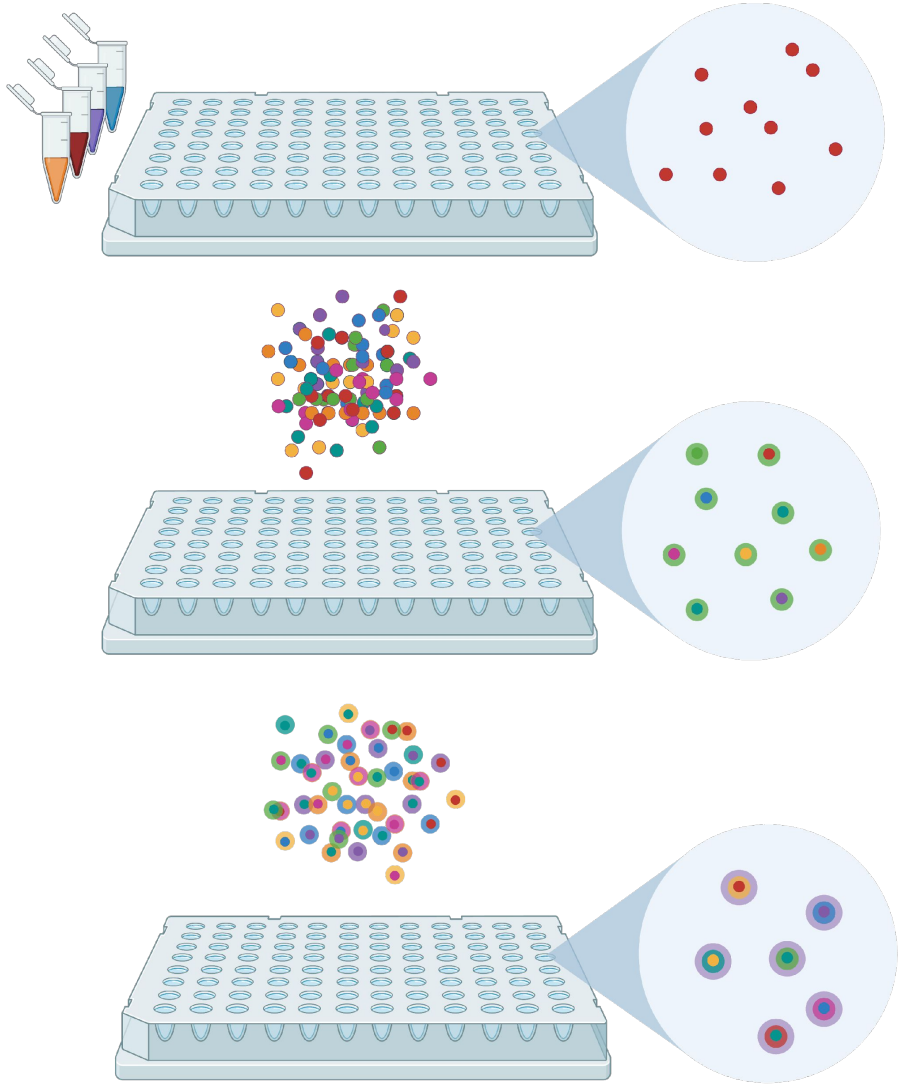

Amplified cDNA processing (dual index)
③
Read 1
UMI
Poly(dT)VN
TSO
10x Barcode
Enzymatic Fragmentation
End Repair, A-tailing, Ligation
Read 2
Cleanup & Priming
Sample Index (i7)
P7
P5
Sample Index (i5)
Sample Index PCR
P5
Read 1
10x Barcode
UMI
Poly(dT)VN


Final library
Sample Index (i5:10)
Read 1:28
10x BC+UMI
Sample Index (i7:10)
P5
TruSeq Read 1
10x Barcode
UMI
Poly(dT)VN
Read 2:90 Insert
TruSeq Read 2
P7

Single-cell/nuclei combinatorial indexing

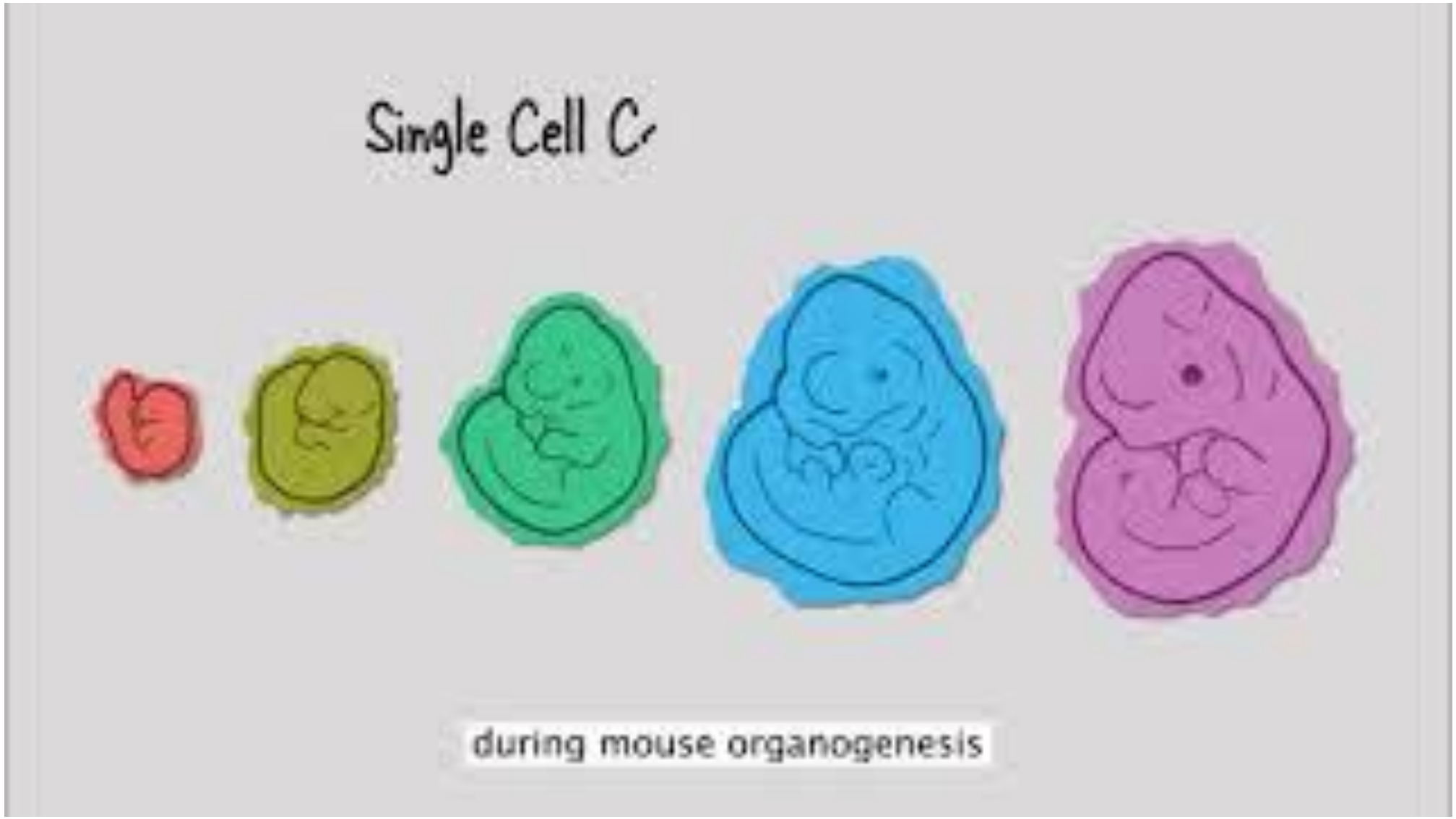
Abbreviated SCI, pronounced “sky”

- Novel, high throughput protocols
- *In situ* molecular indexing and a ‘split and pool’ framework
- Can profile millions of cells in a single experiment
- Scales sub-linearly with cost!

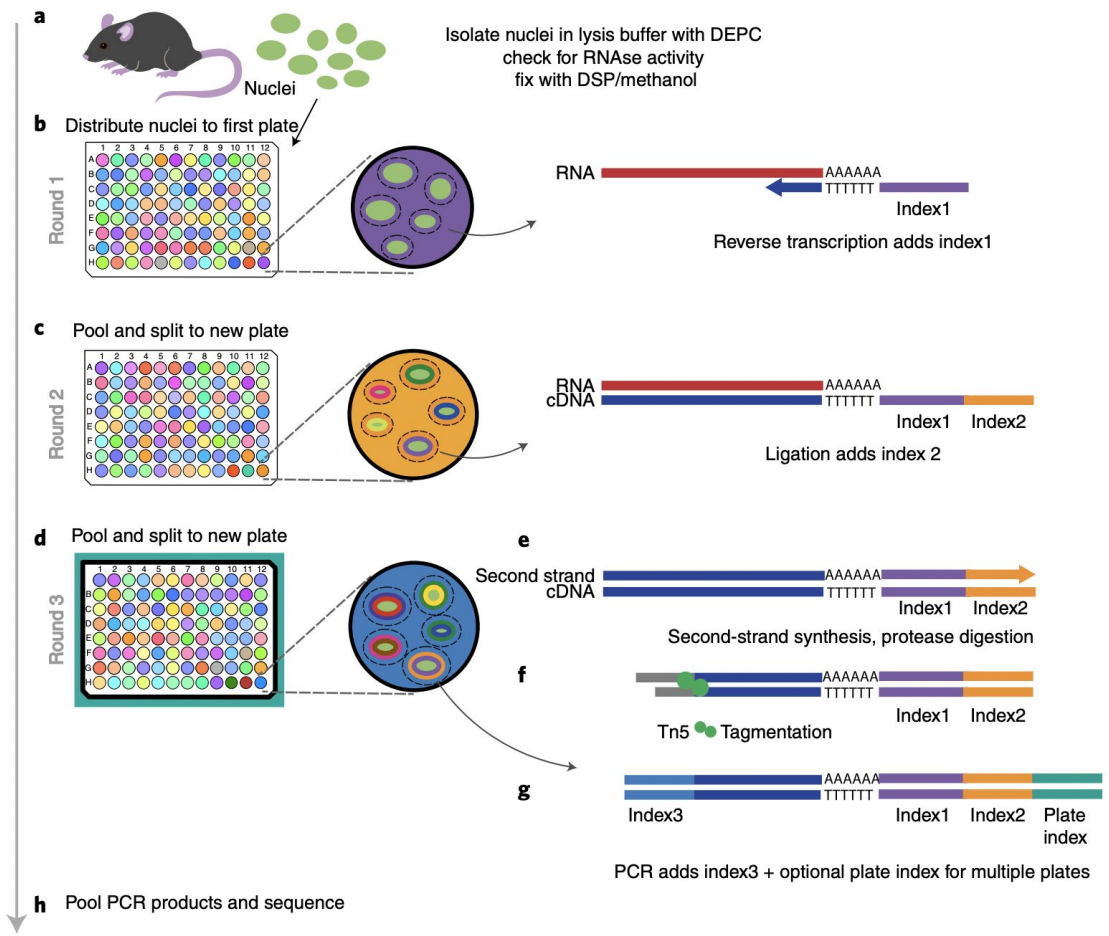


https://youtu.be/VOiCvIujYd8

Single-cell/nuclei combinatorial indexing



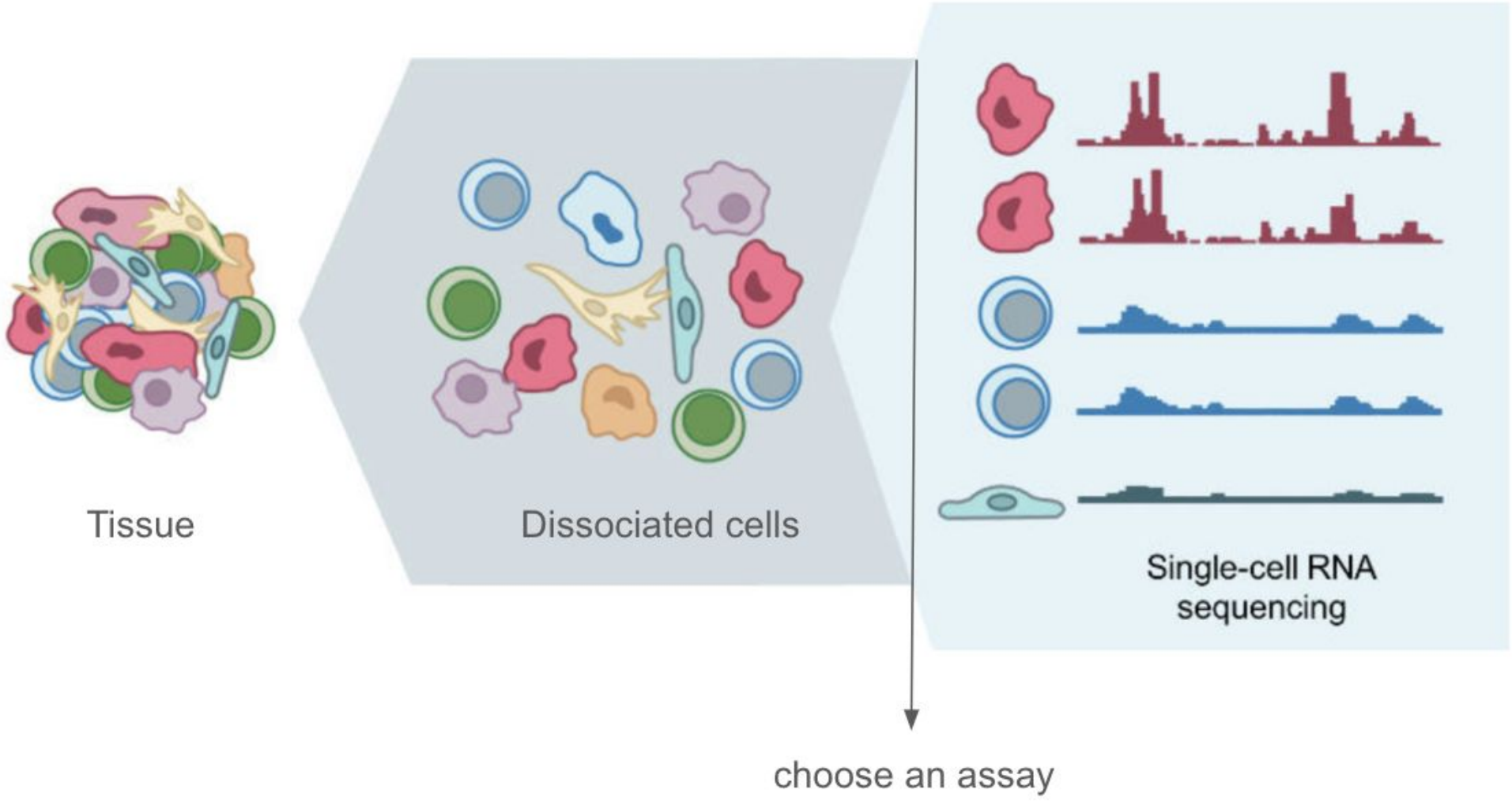
sci-RNA-seq3



- Cells/nuclei are distributed across wells.
- Each well contains a unique barcode. First-strand synthesis/ reverse transcription (RT) labels all cells in the well with the barcode.
- Cells are pooled and randomly spit into wells, and a second barcode is added.
- Exponentially scaled through increasing numbers of barcoding rounds!

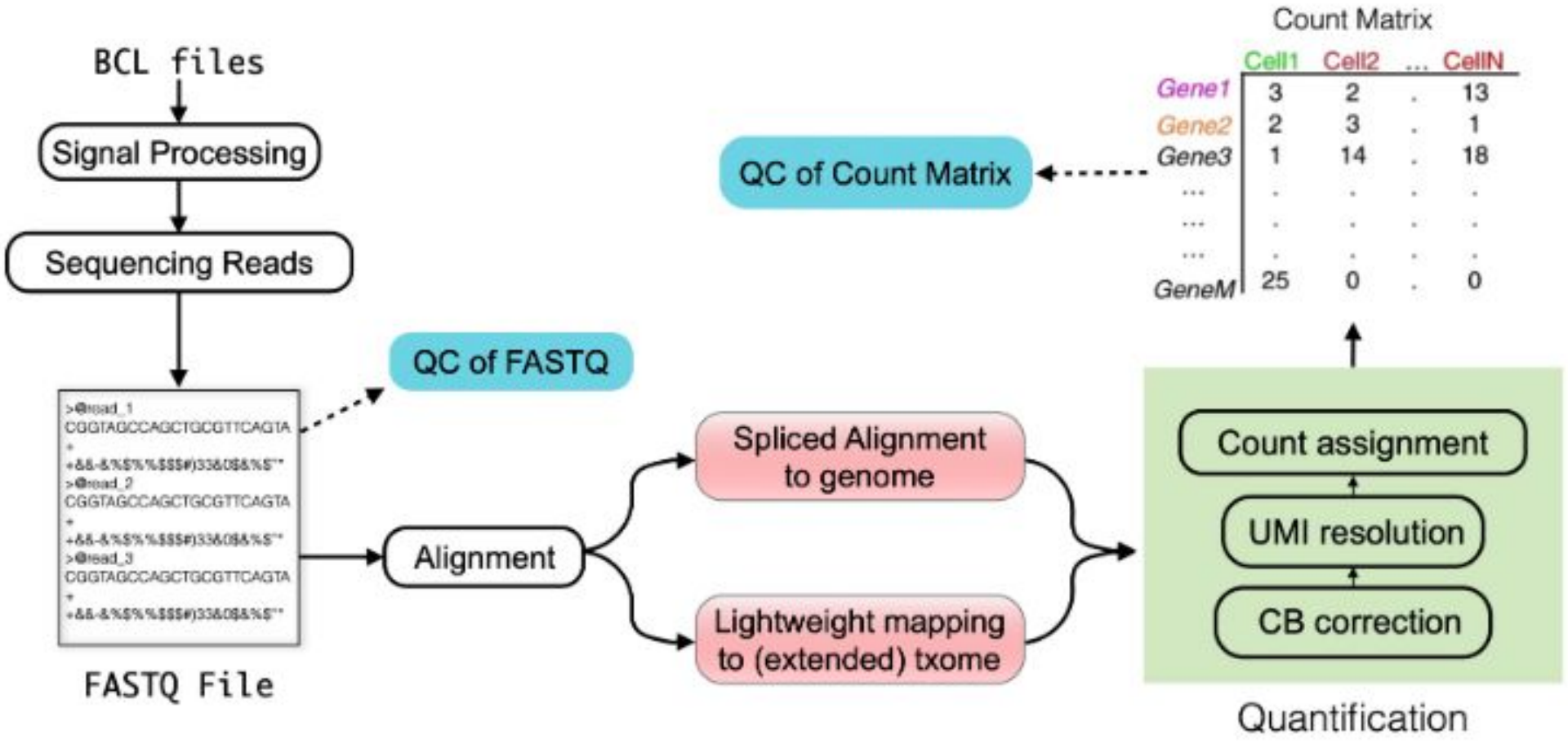
Martin et al. 2023 (Nature Protocols)

Recap



Modified from Itay Tirosh & Mario Suva

Next time: Analysis Overview



https://www.sc-best-practices.org/