

Journal Club: Single Cell Technologies

Hanqing Liu (Ecker Lab)

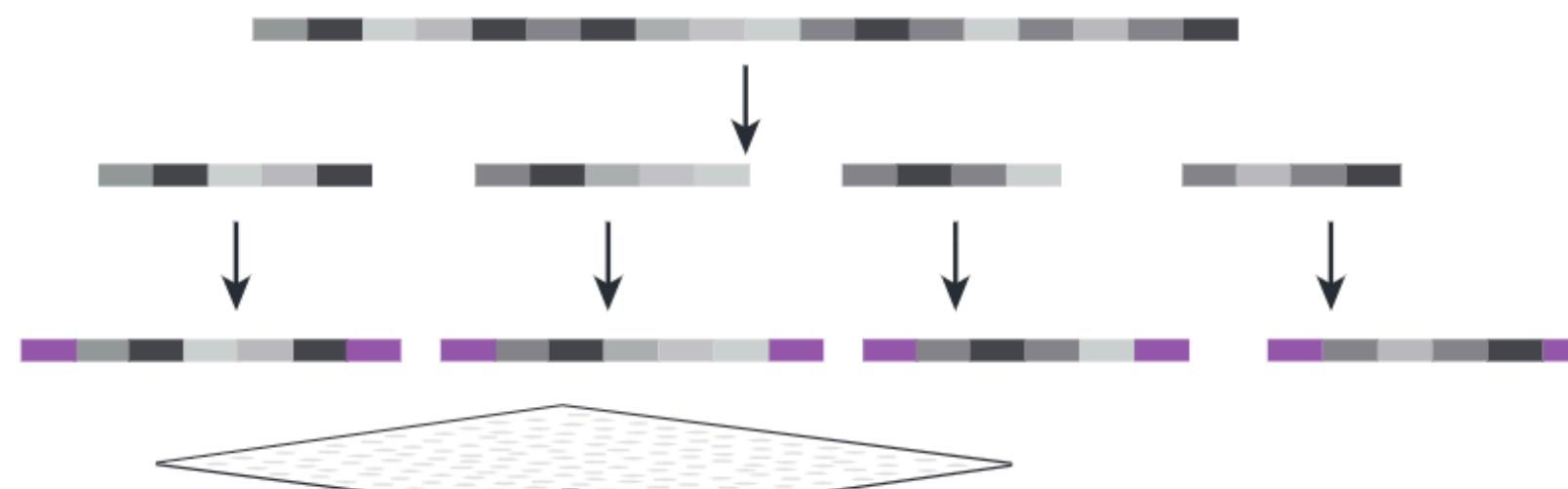
Howard Hughes Medical Institute
The Salk Institute for Biological Studies
UCSD Biology Graduate Program

Introduction

DNA Sequencing Technologies

Second generation sequencing (massively parallel)

1 Genomic DNA



2 Fragmented DNA

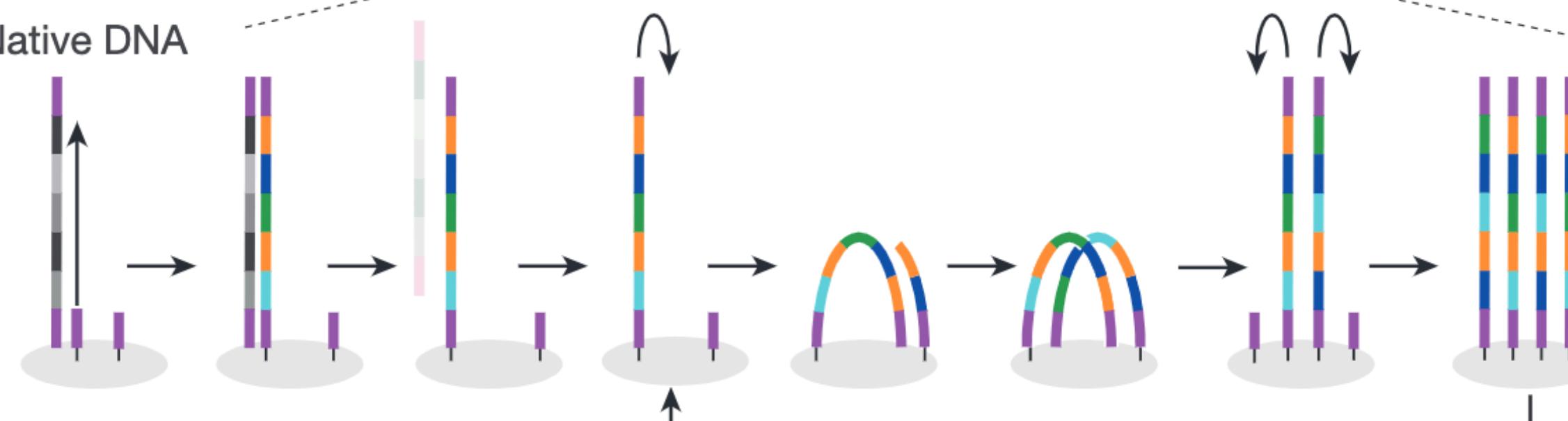


3 Adaptor ligation

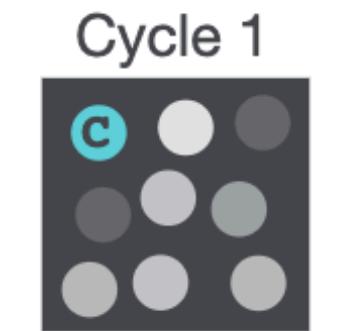


4 Amplification

Native DNA



5 Detection



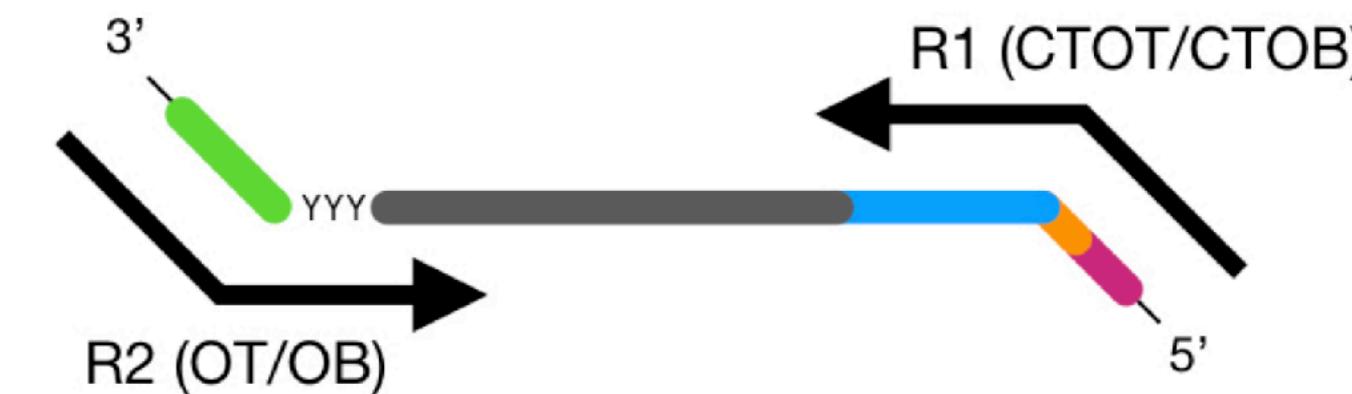
3' ... G A C T A G A T C C G A G C G T G A ... 5'
5' ... C T G A ...

Shendure et al., (2017) Nature

- All sequencing technologies are eventually DNA sequencing on illumina sequencer.

- Three questions to understand a sequencing technology:

1. Which part of the molecular information does it profile?
2. How to turn those molecular information into DNA fragments with **adapters**? (i.e. Library Preparation)
3. How to **multiplex** sample to increase the capacity of single sequencing batch? (i.e. including barcodes in the DNA fragments?)

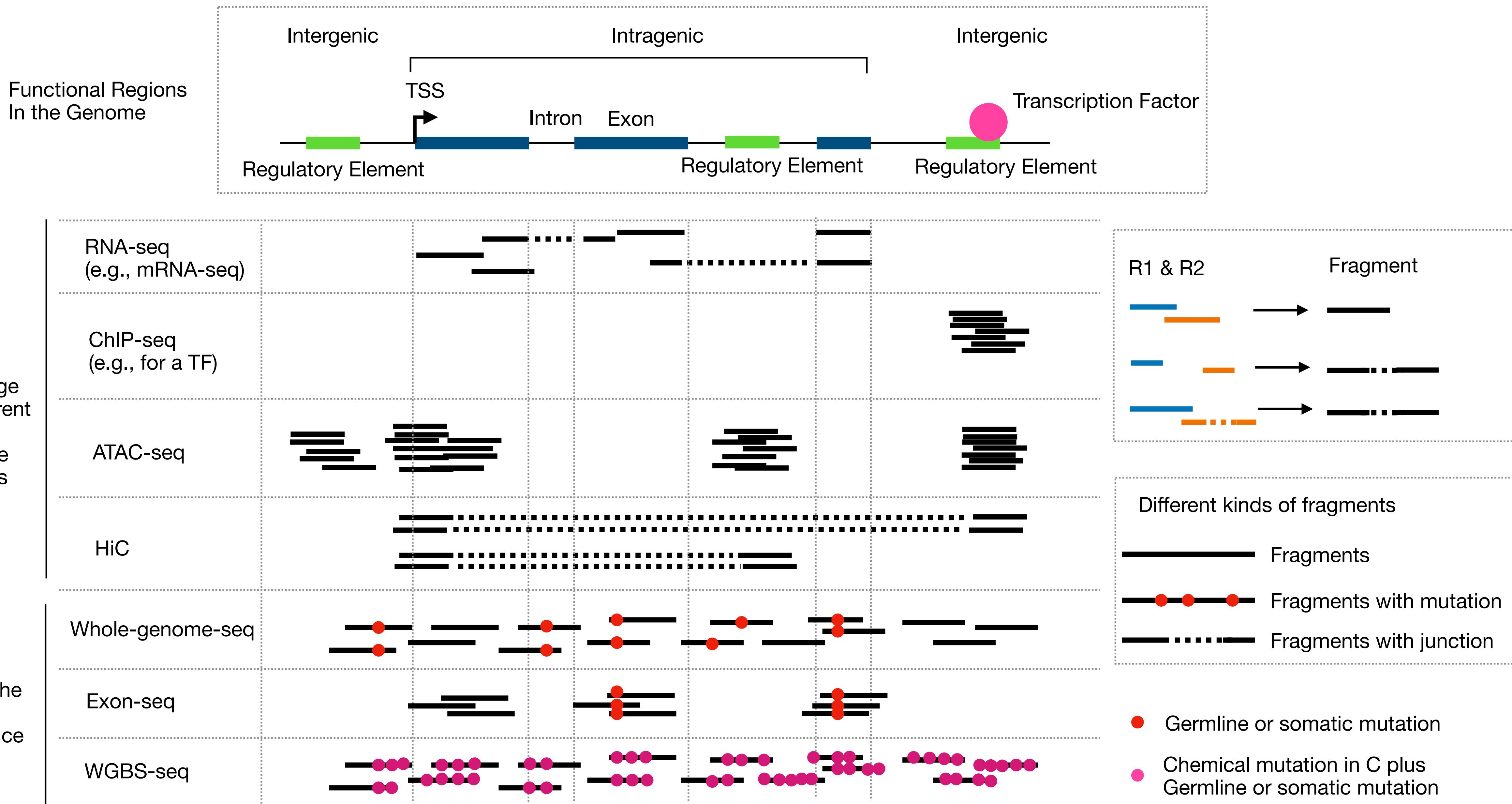


- Post-Bisulfite Genomic DNA (CT)
 - RN-H Random Primer Part
 - Hexamer Index
 - P5 Primer Part
 - P7 Primer Part
- Y=C, T, added by adaptase

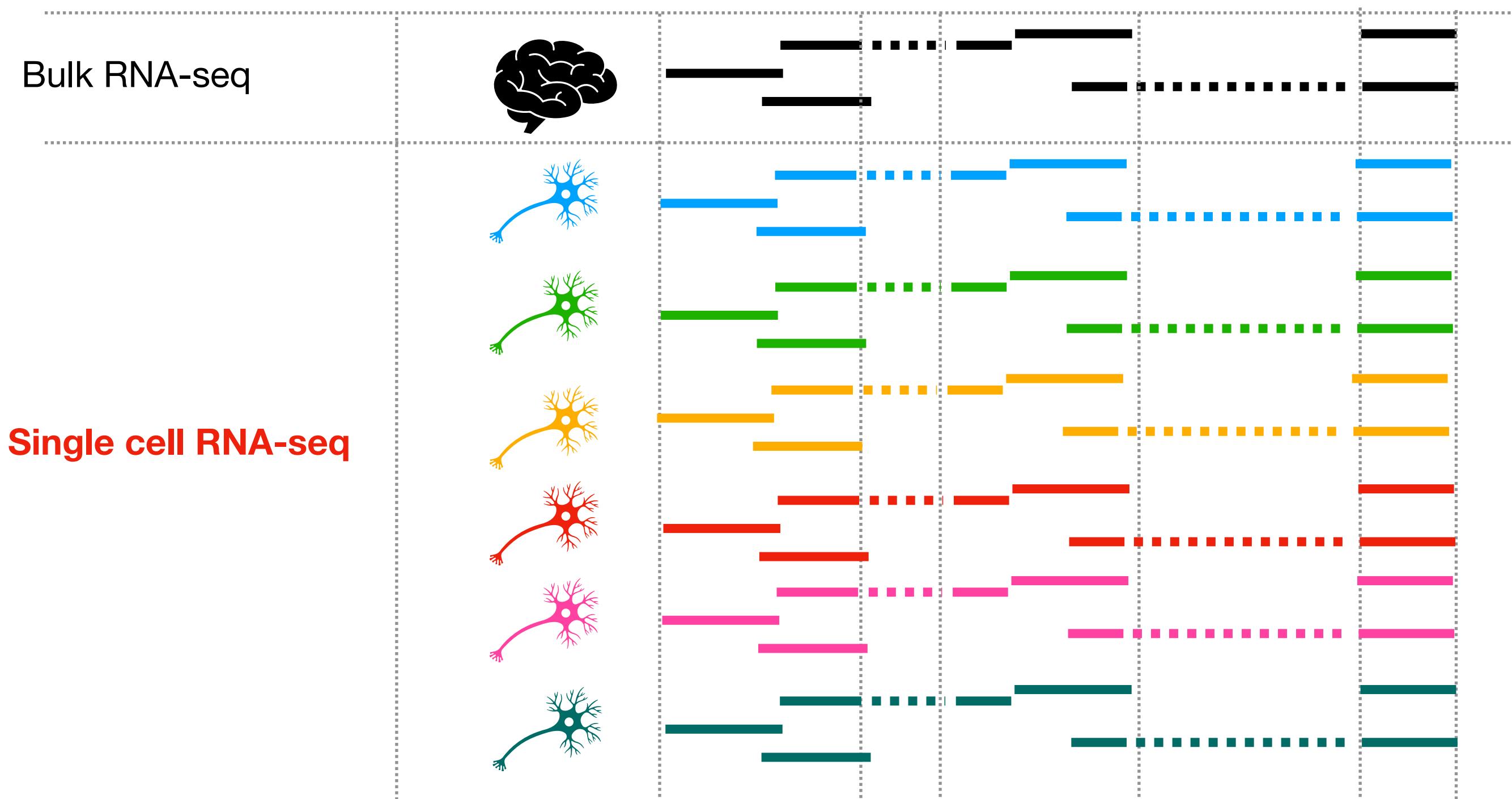
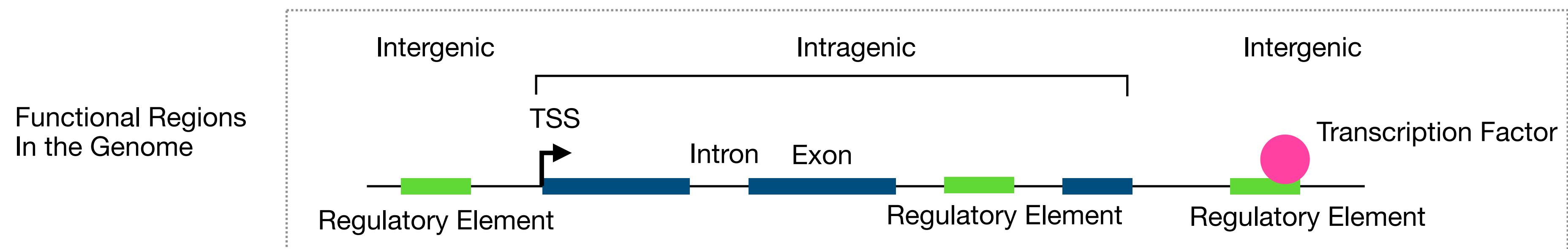
Input fragments structure in snmC-seq2

Sequencing technologies are all about how to prepare these DNA fragments.

Sequencing Technologies - Classified by the Type of Quantification



From Bulk-Seq to Single-Cell-Seq



Bulk



Single Cell



Contents

- Four key questions to understand most single cell technologies
- Explain common solutions to the single-cell specific questions
- Future direction of single-cell technologies

Questions To Understand Single Cell Technology

Technology Basics

1. How cells are isolated?

2. How cells are barcoded?

3. What is the library preparing protocol?

4. Which molecular modality/modalities does it profile?

Q1, Q2 is single cell specific, Q3, Q4 is similar to the bulk-seq, its just that the reaction usually happens in a cell-basis

Other consideration

1. Cellular throughput?

2. Cellular Sensitivity? (Input cell vs final cell being detected)

3. Coverage / Molecular Sensitivity? (gene per cell or reads per cell)

4. Cost?

5. Special sample requirement?

6. Easy to adopt?

How Cells Are Isolated?

1. Plate-based - FACS to 96/384 plate
 - medium
 - high throughput - via combinatorial index
2. Plate-based - Micro-well plate
 - high throughput
3. Bead-based with microfluidic equipment
 - high throughput

Cellular throughput

- Medium (100 - 1000)
- High (1000 - 10000)

Plate Based Technology - FACS/FANS To Plate - Medium Throughput

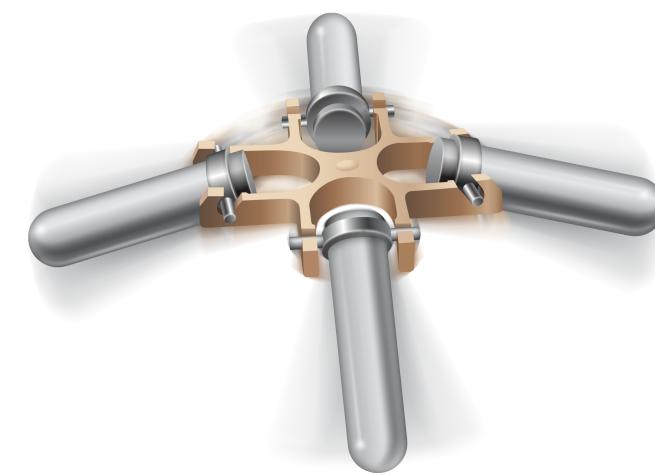
snmC-seq2 as an example, other plate based tech (SMART-seq for RNA) are pretty similar



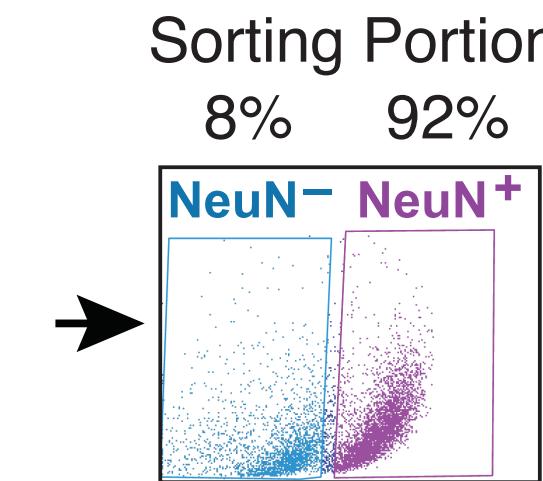
Tissue preparation

fresh or frozen tissue?

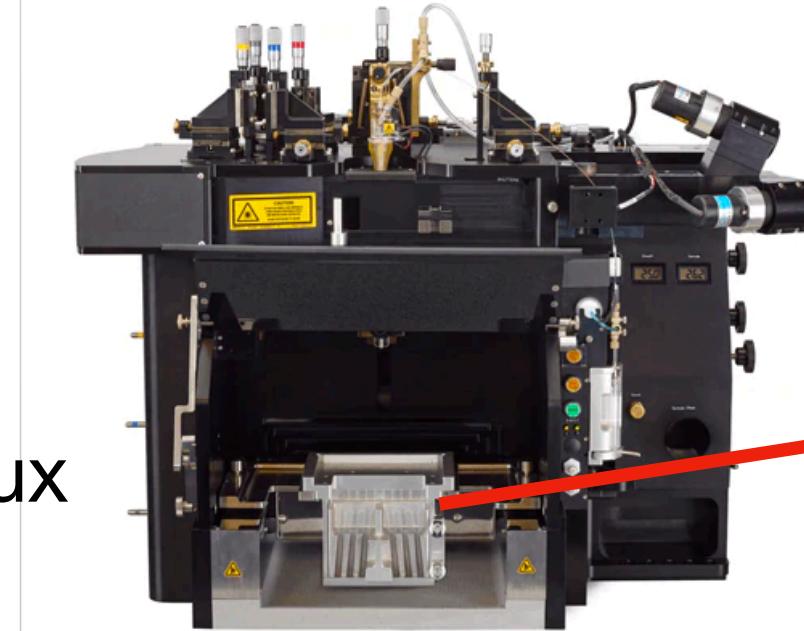
If human, how long is the post-mortem interval?



Single cell/nuclei preparation



BD Influx



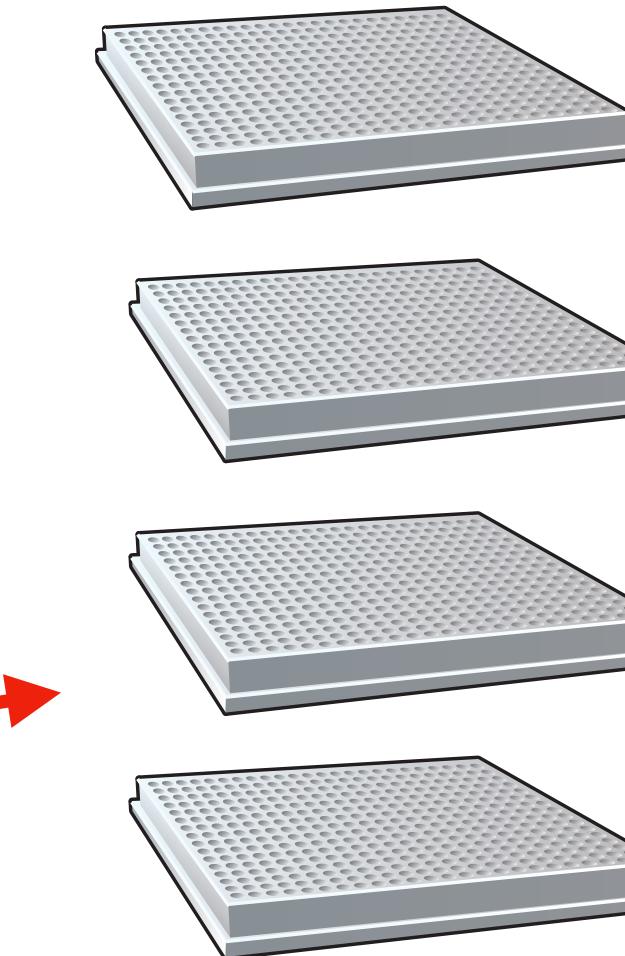
FACS / FANS

Use antibody to select sub-population?

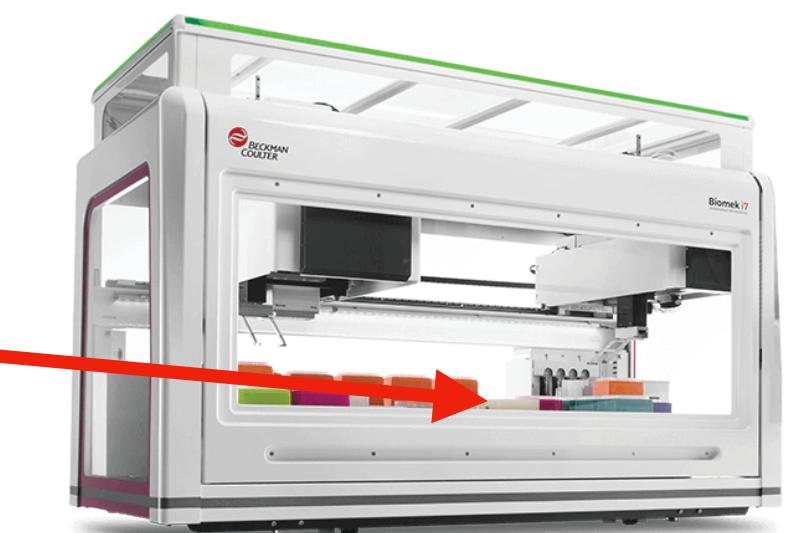
Sorting speed?

Cell efficiency?

Doublents?



One well one cell one reaction



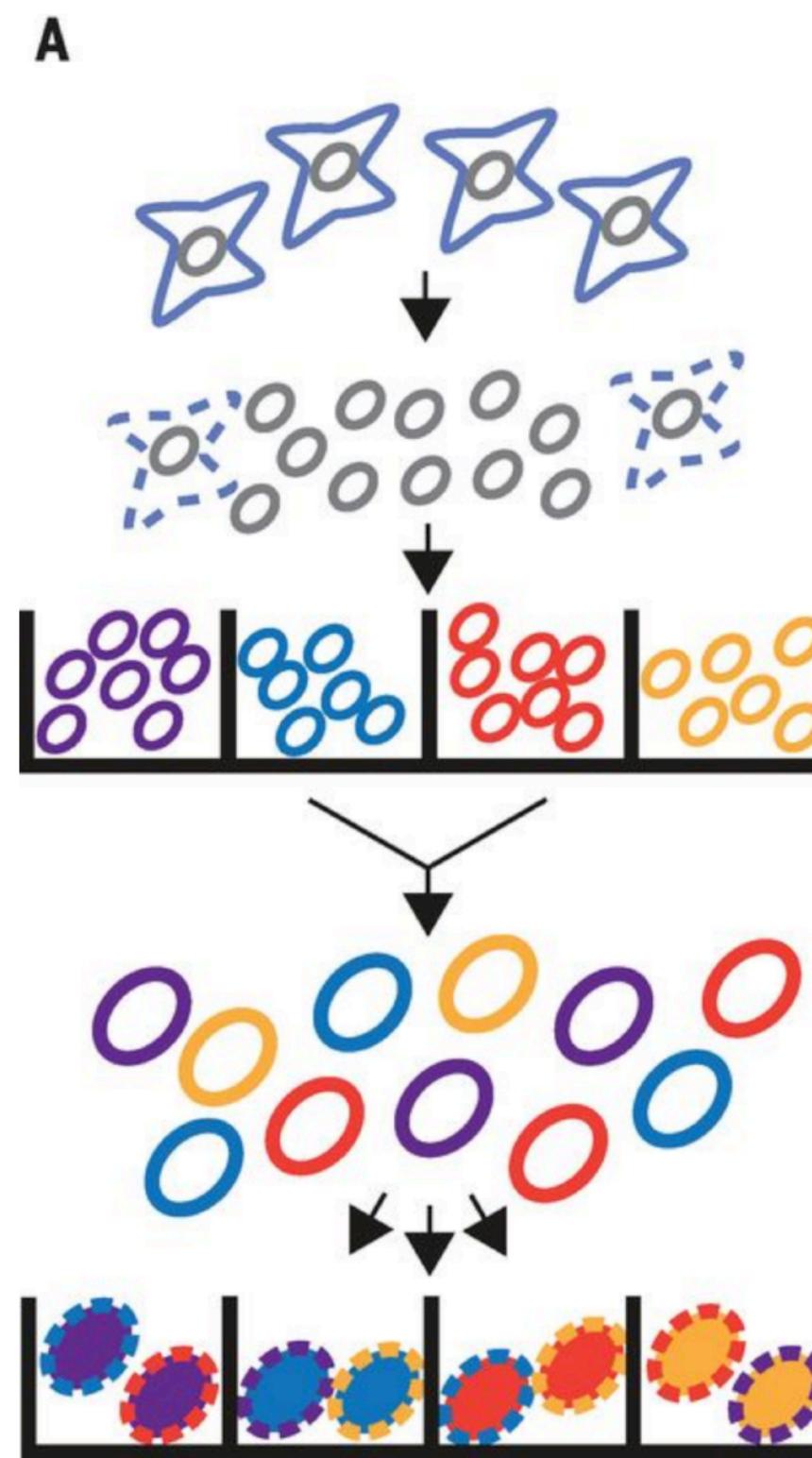
Beckman i7

Library prep w/ liquid handler

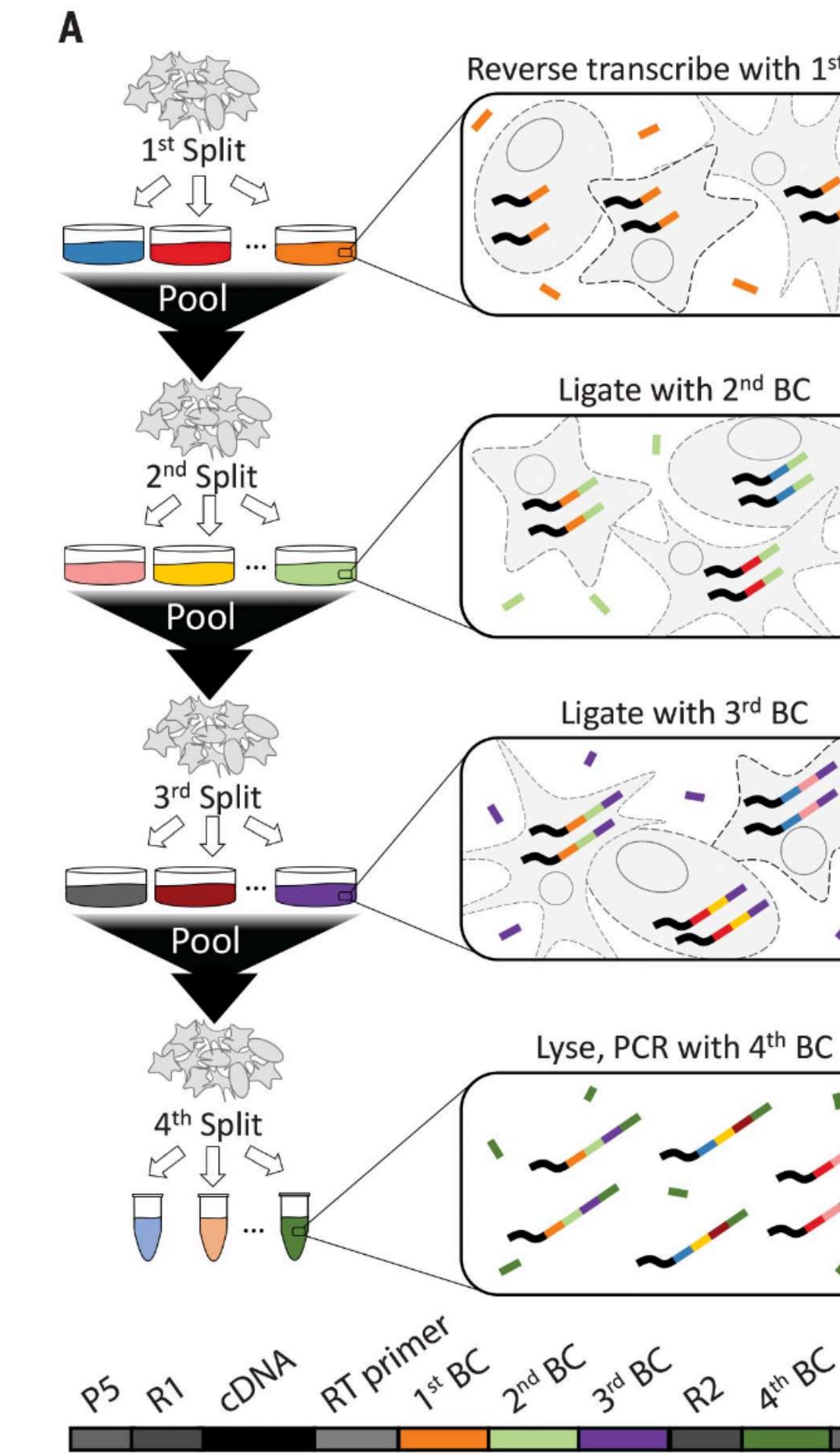
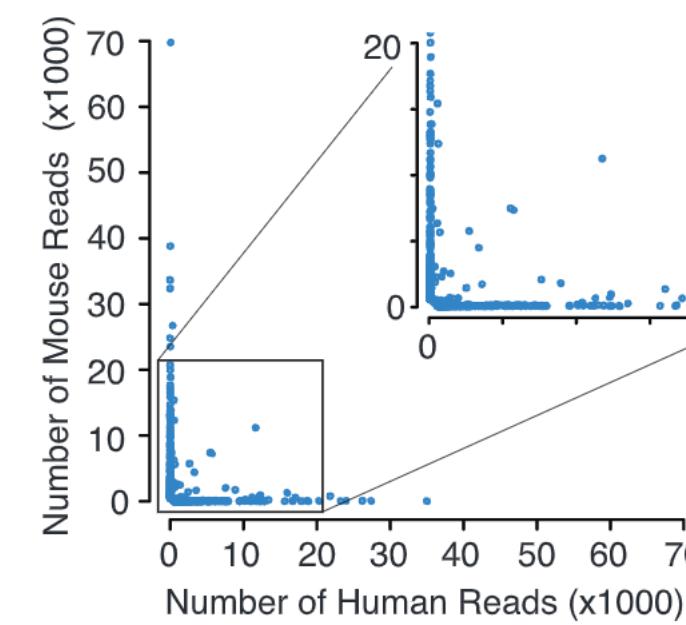
Needs time and efforts to setup and specialized person to handle

Plate Based Technology - FACS/FANS To Plate - High Throughput

Combinatorial cellular index



Human-cell mixture experiment to test doublet ratio.



Rosenberg 2018 science

Multiple cell per well

Split - Pool strategy +
Combinatorial index

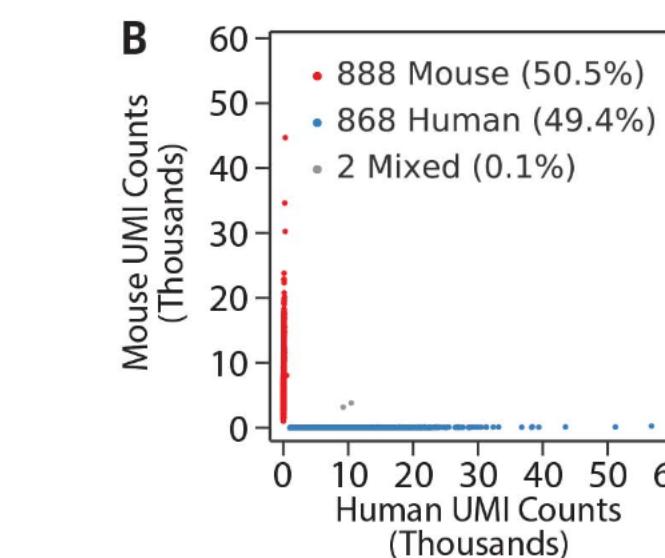
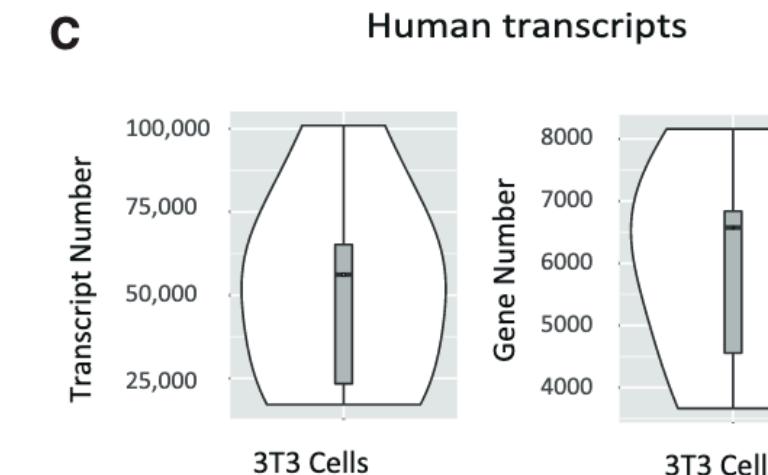
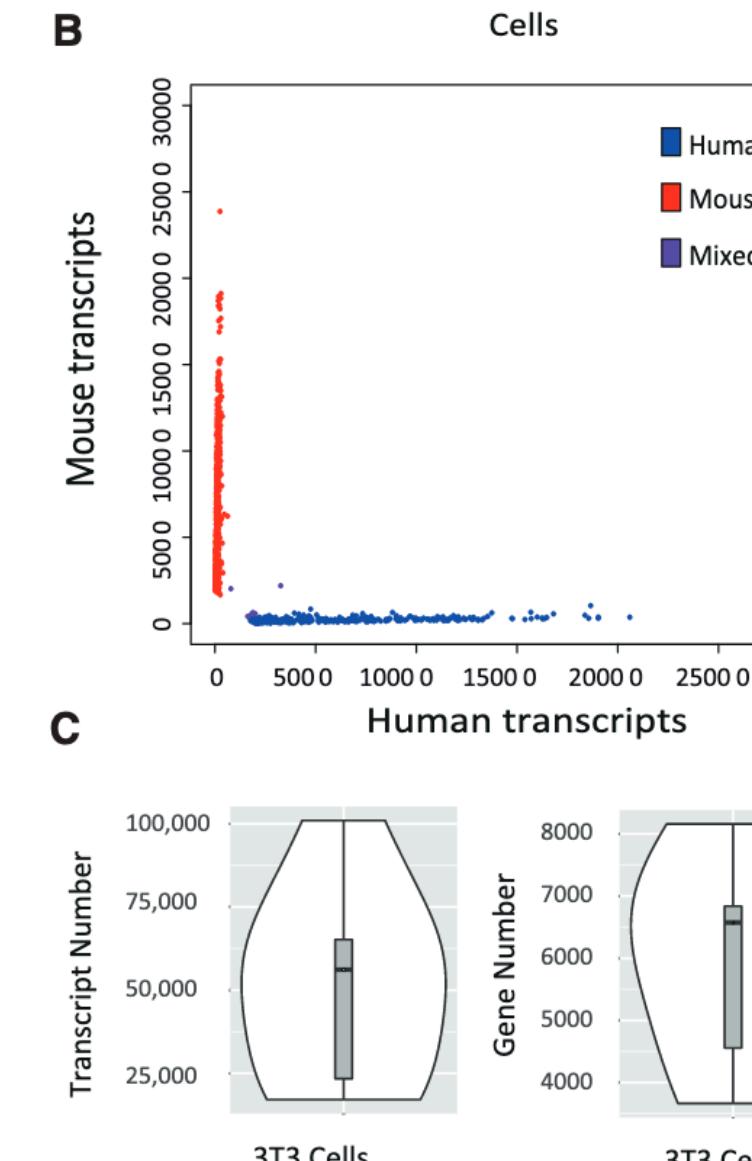
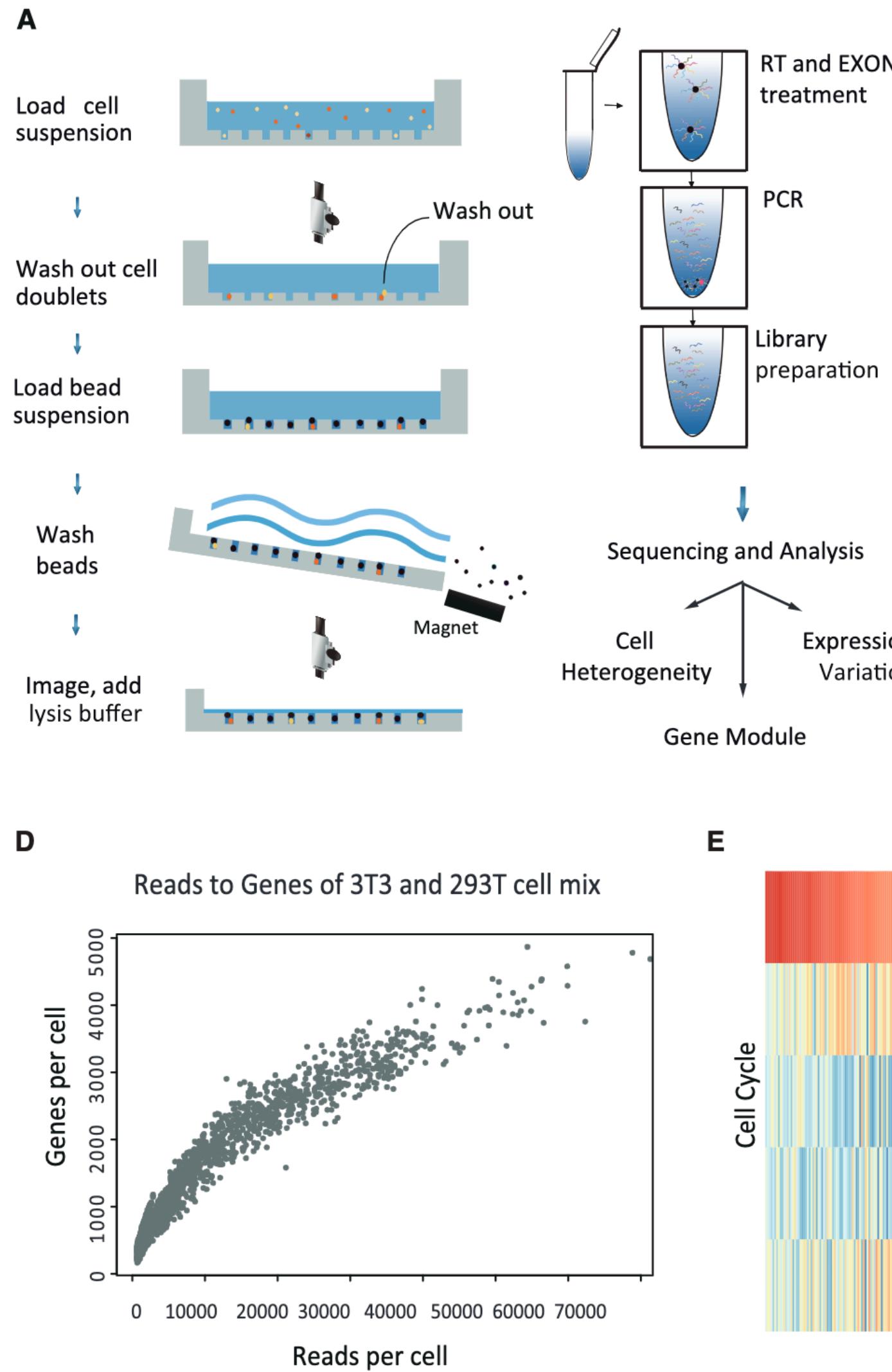
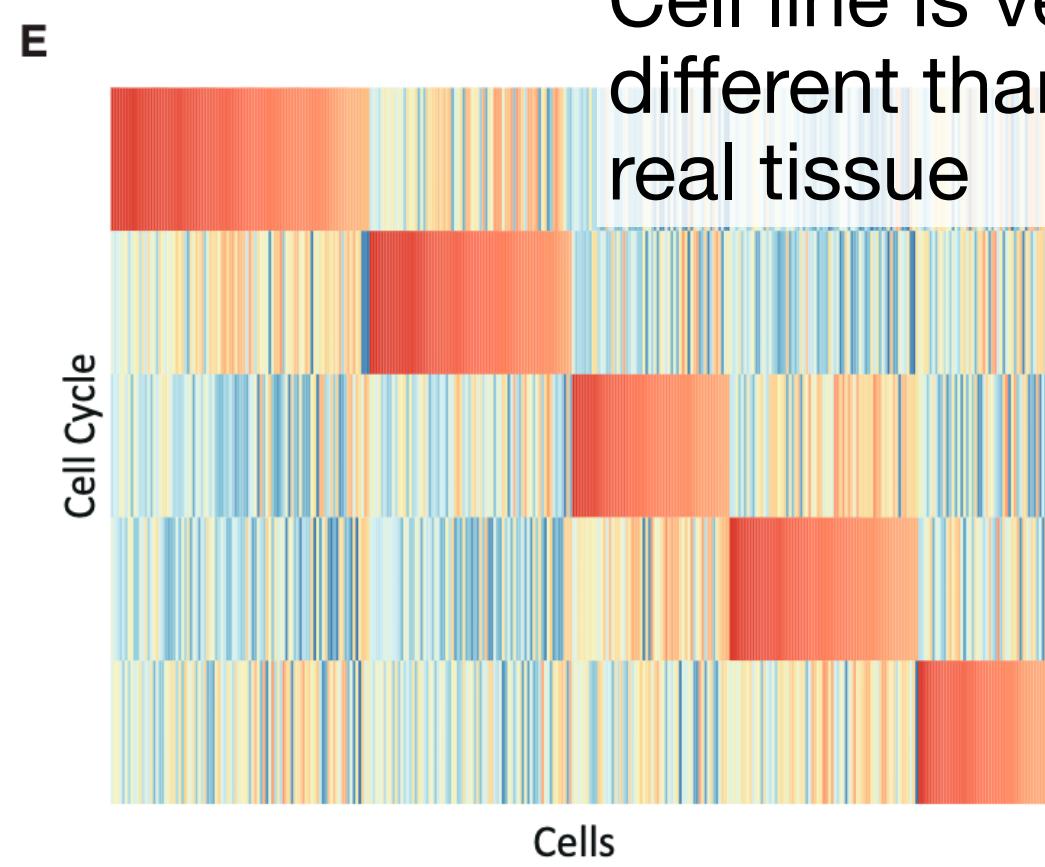
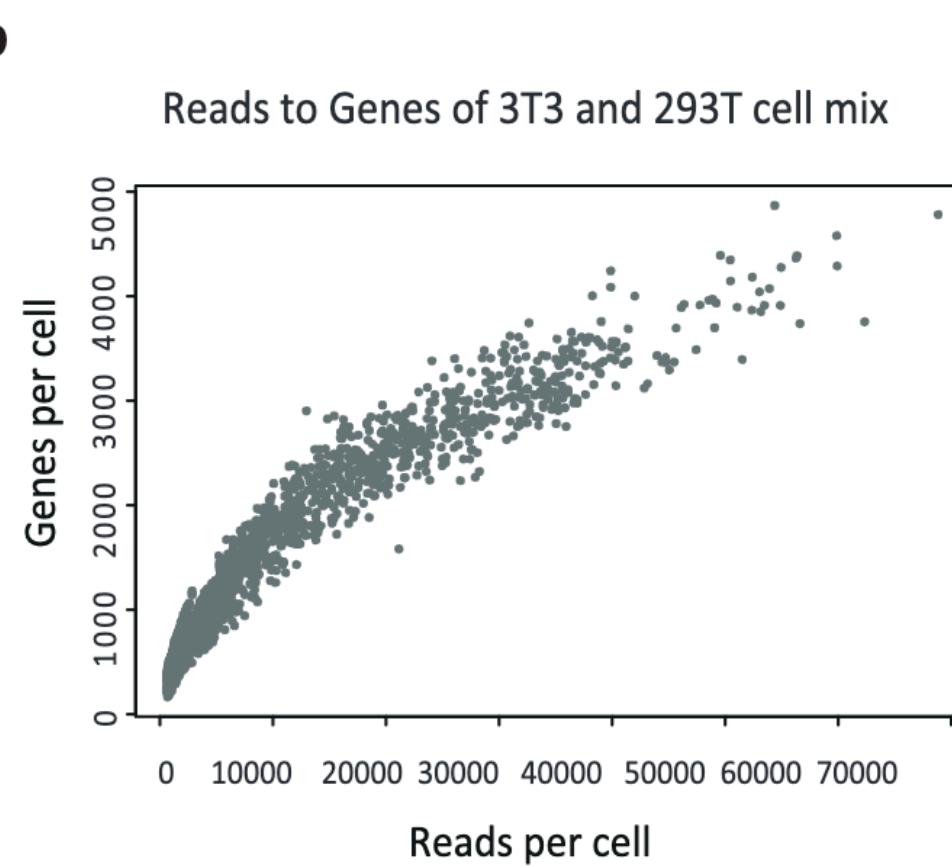


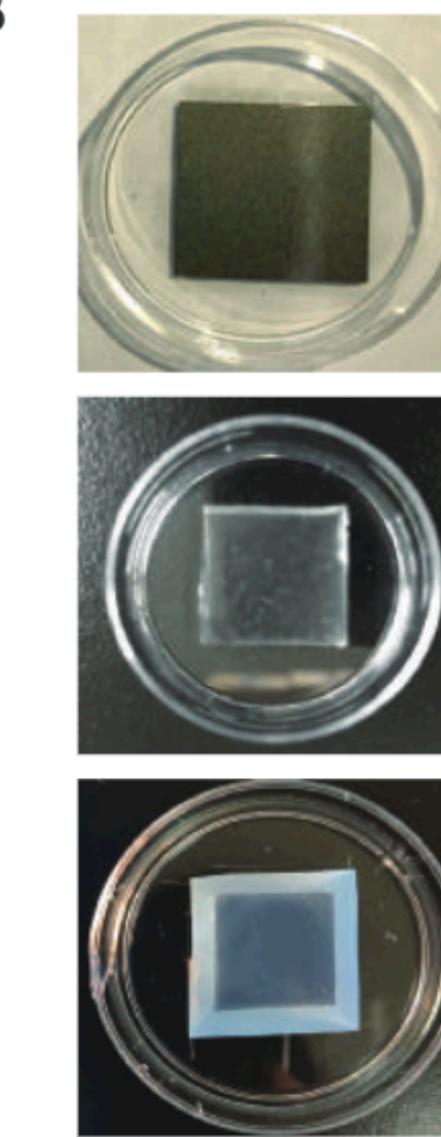
Plate Based Technology - Micro Well Based



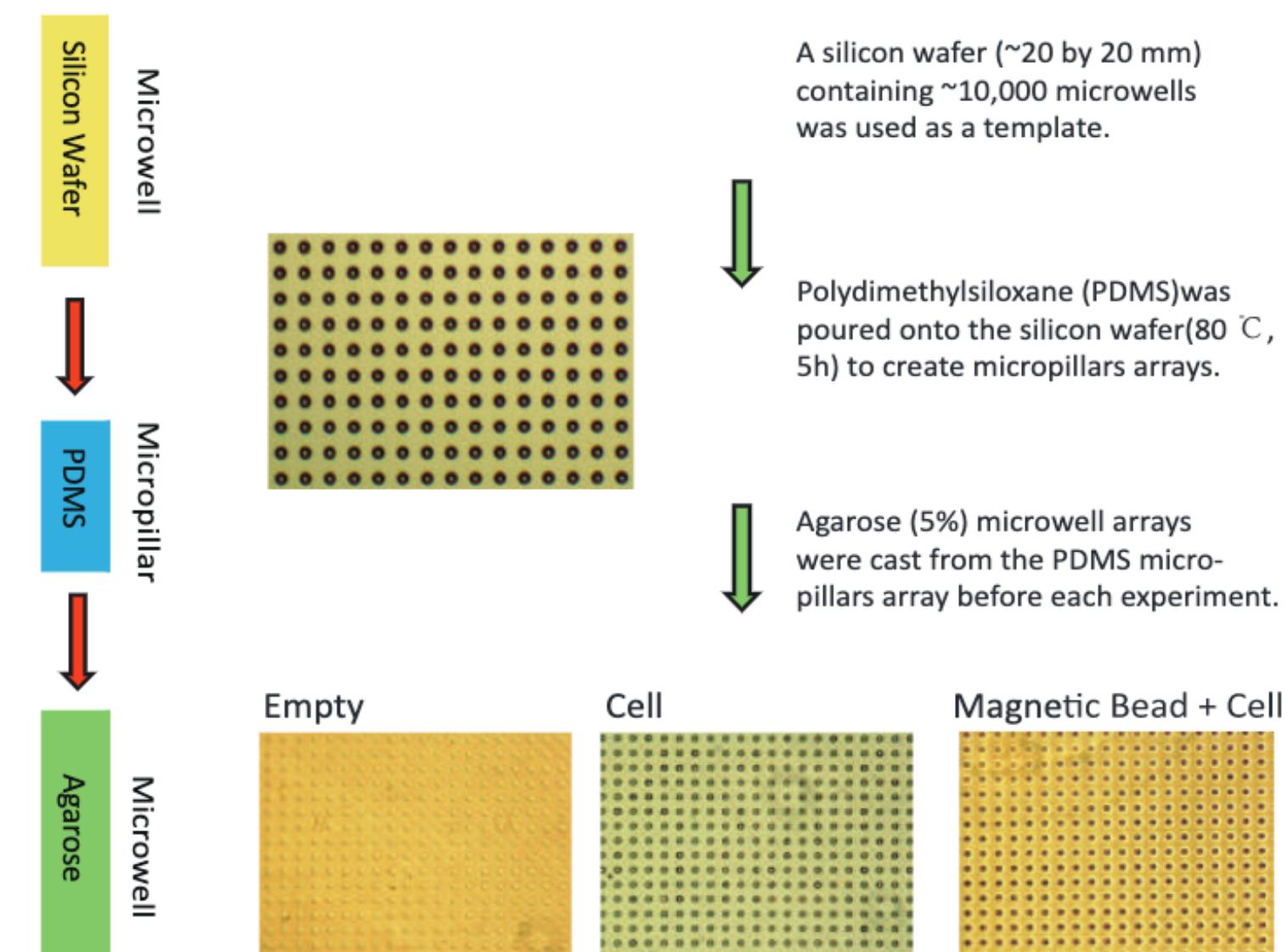
Cell line is very different than real tissue



B

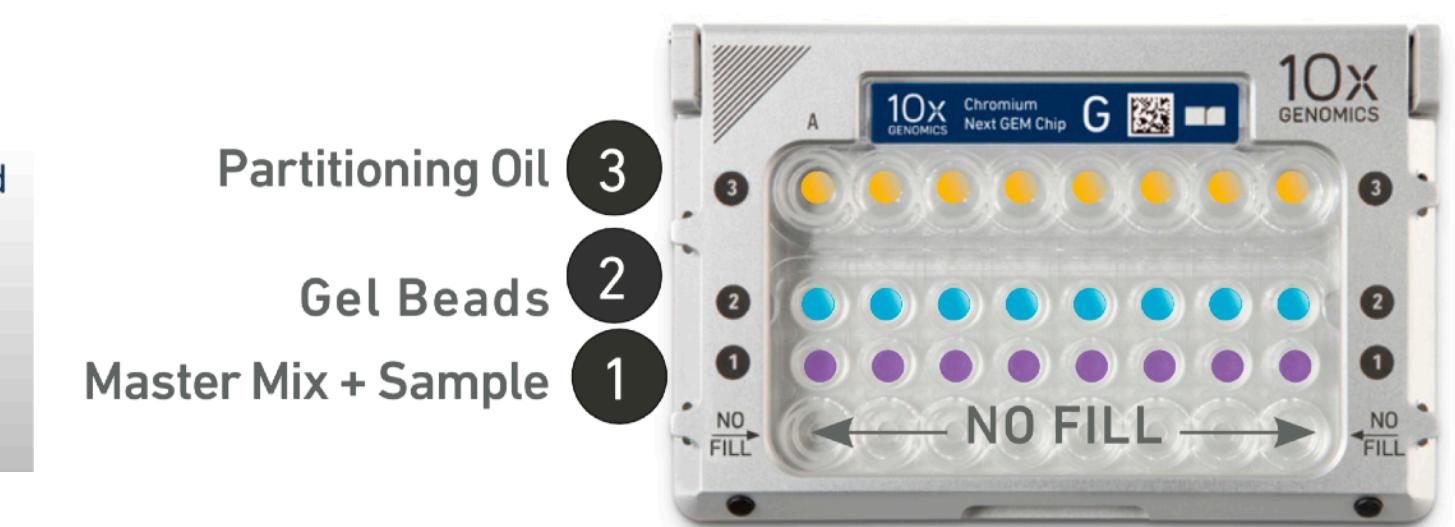
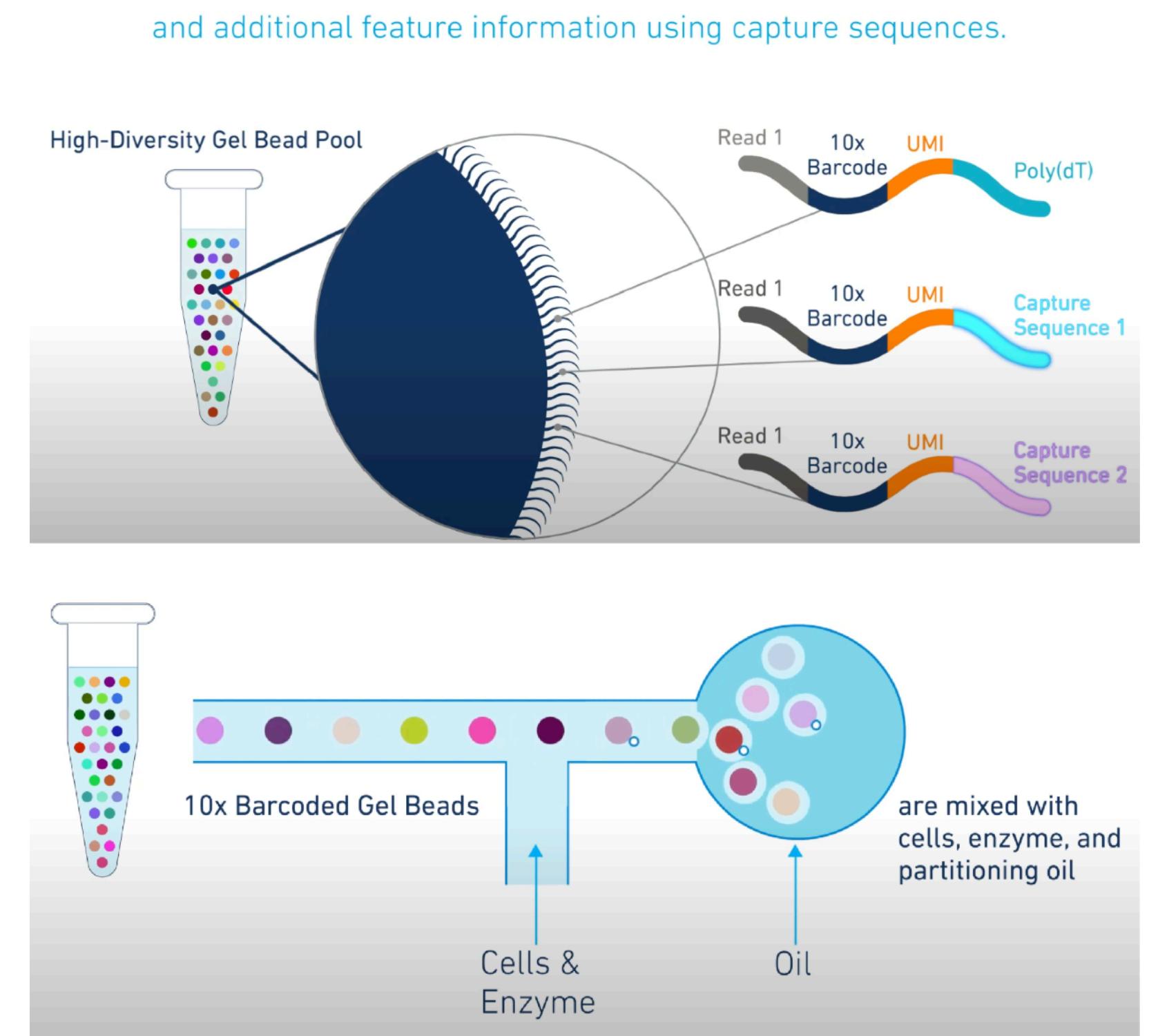
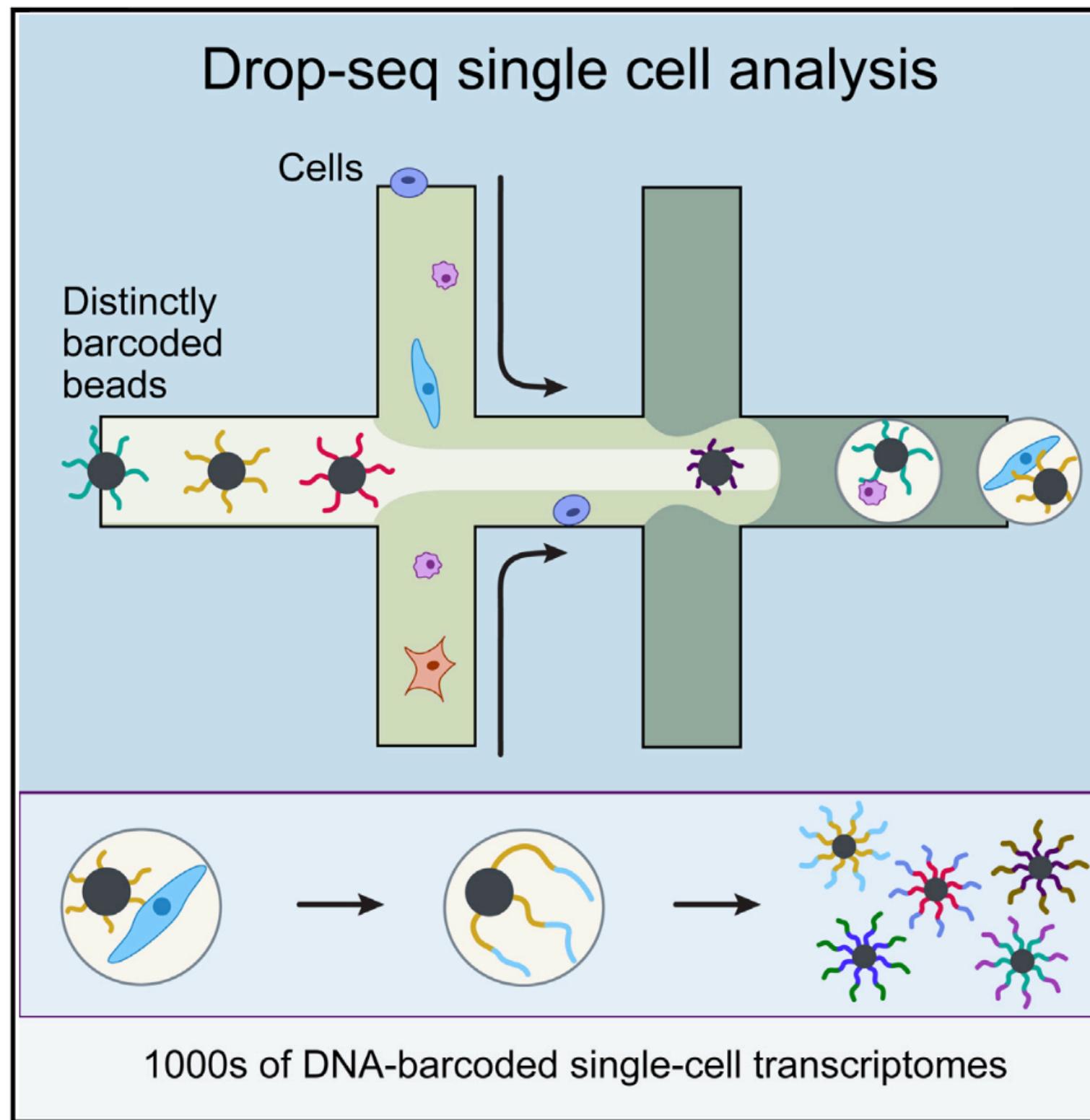


Cheap



high cellular throughput

Bead-Based With Microfluidic Equipment



Drop-seq
Macosko 2015 cell

10X genomics

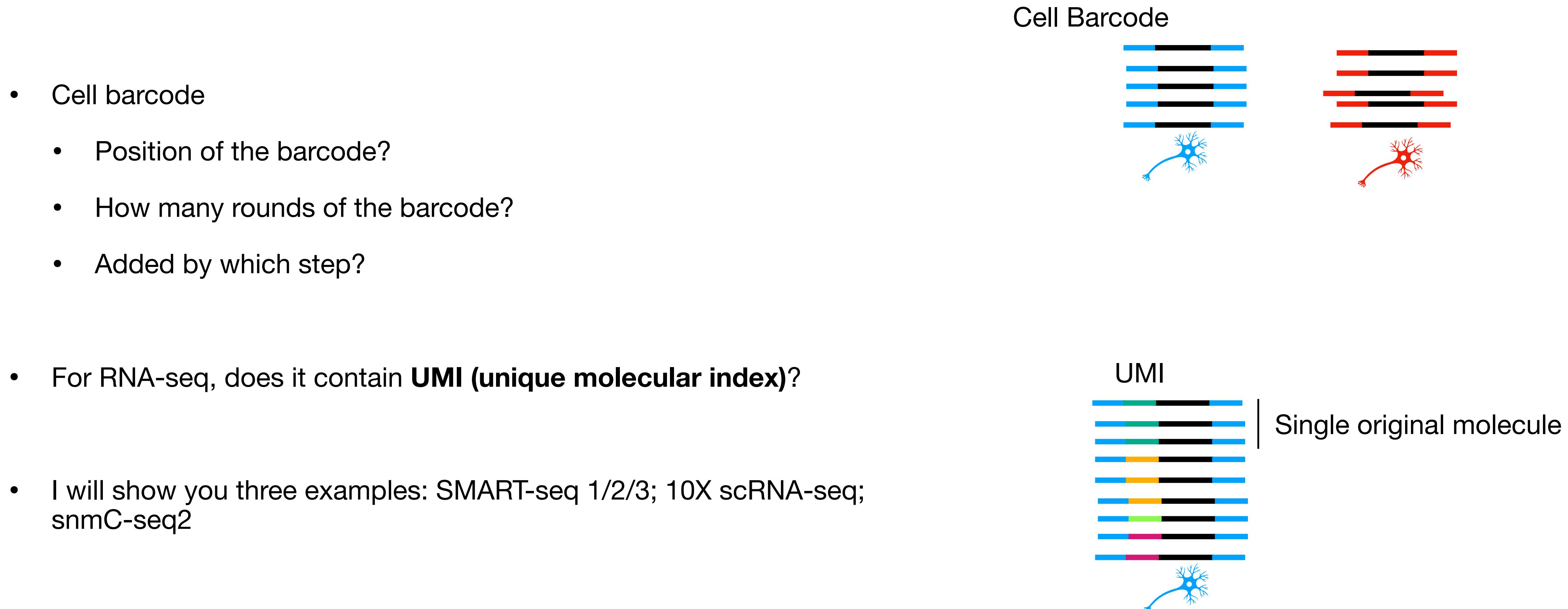
How Cells Are Isolated?

1. Plate-based - FACS to 96/384 plate
 - medium throughput: snmC-seq, SMART-seq,
 - high throughput: snATAC-seq (combinatorial index), SPLiT-seq (combinatorial index, RNA)
2. Plate-based - Micro-well plate
 - Microwell-seq, SeqWell
3. Bead-based with microfluidic equipment
 - 10X scRNA, 10X scATAC, Drop-seq

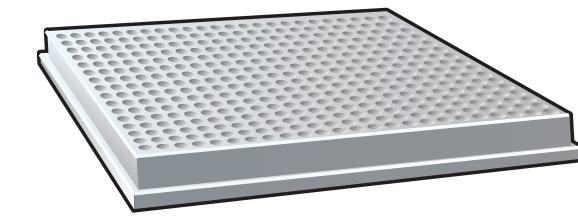
Typically, high cellular throughput means low molecular sensitivity and/or low cellular sensitivity. Low cost per cell.

How Cells Are Barcoded?

- DNA fragments from each cell must have unique barcode before pooling together and put on sequencer



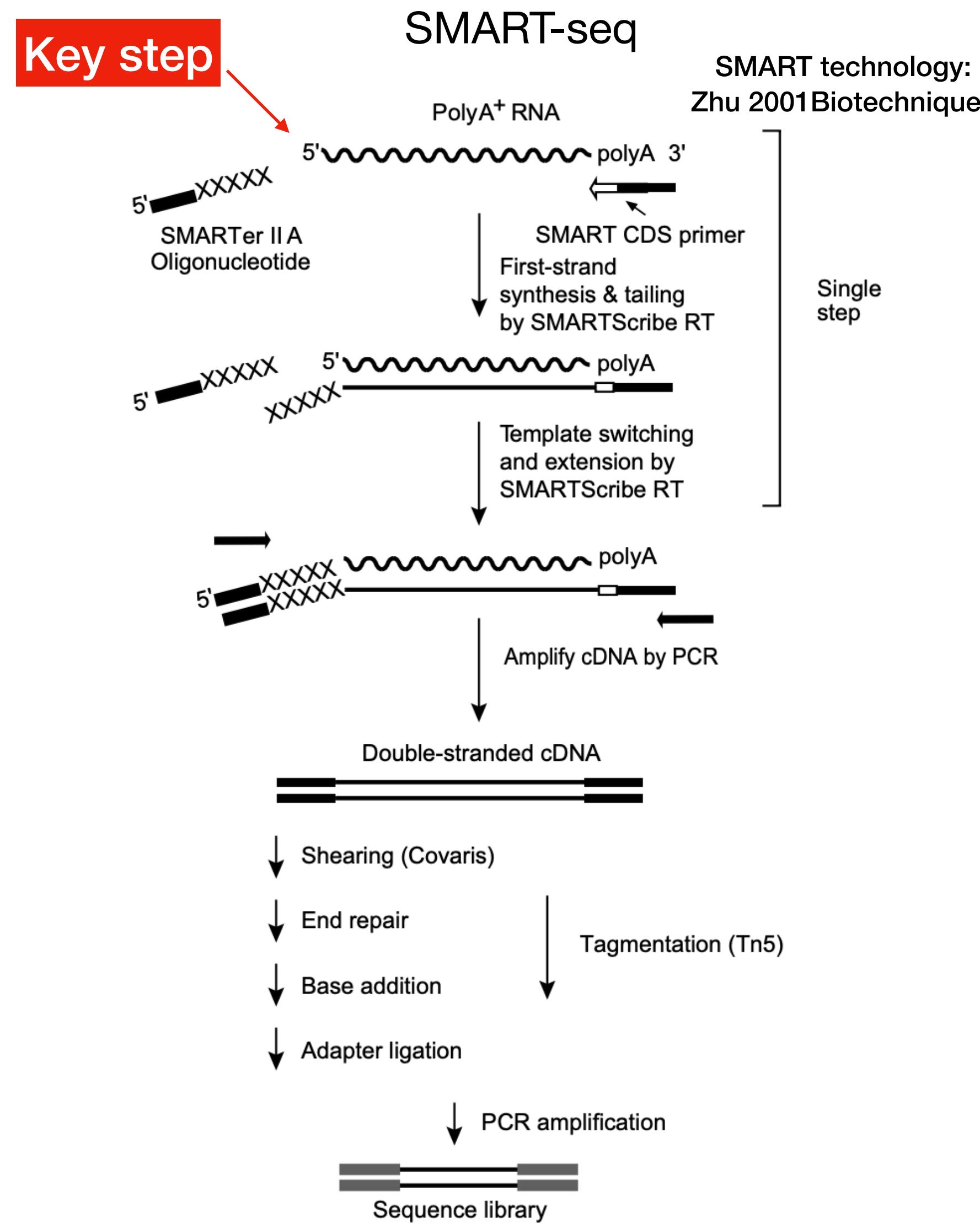
SMART-Seq 1, 2, 3 - Key: Full Length cDNA, Isoform, Medium Throughput With High Coverage



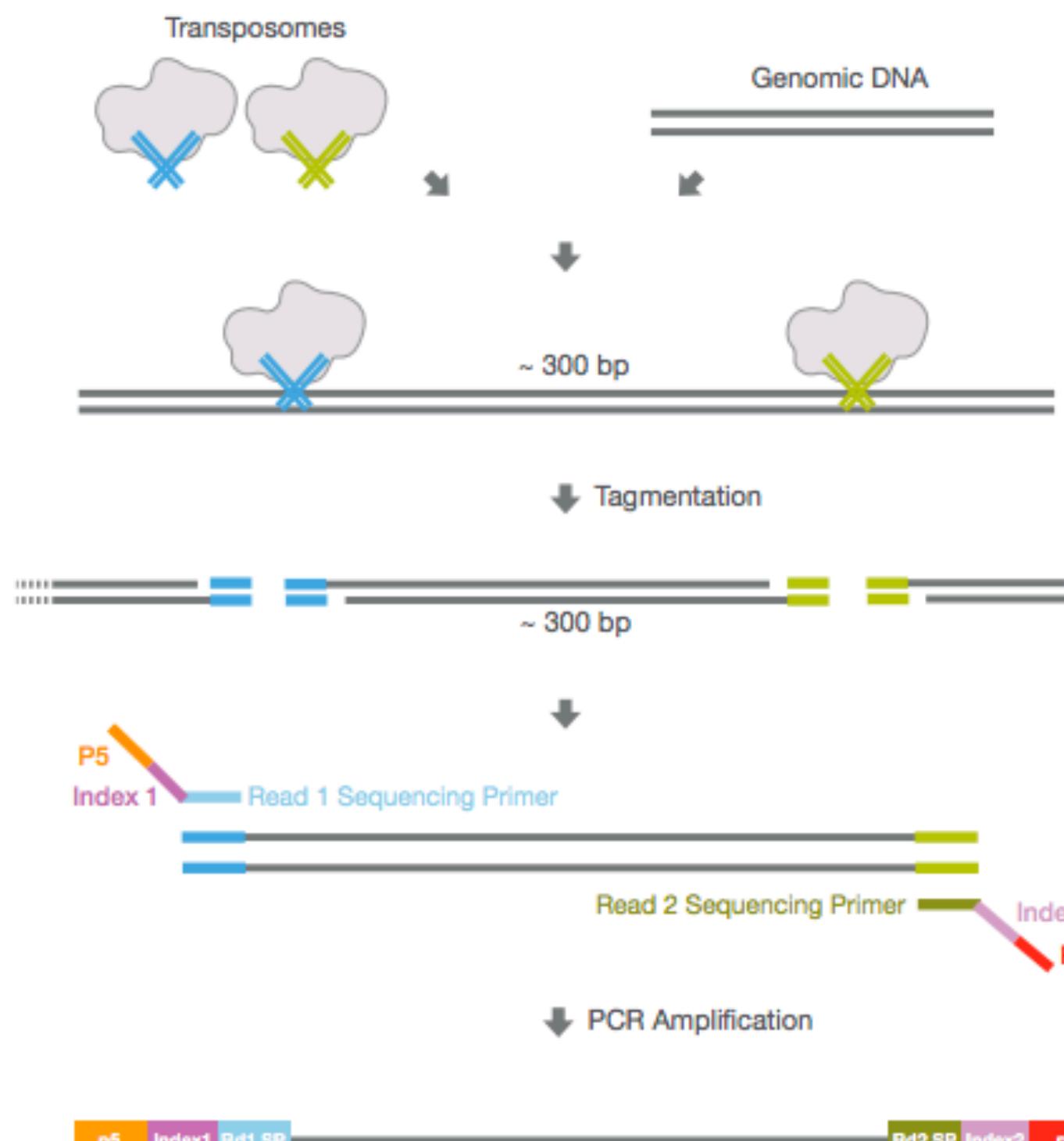
One well one cell
one reaction

Remember SMART-seq is plate-based,
all following steps happen in a well in the plate

SMART-Seq 1, 2, 3 - Key: Full Length cDNA, Isoform (and Allele), Medium Throughput With High Coverage

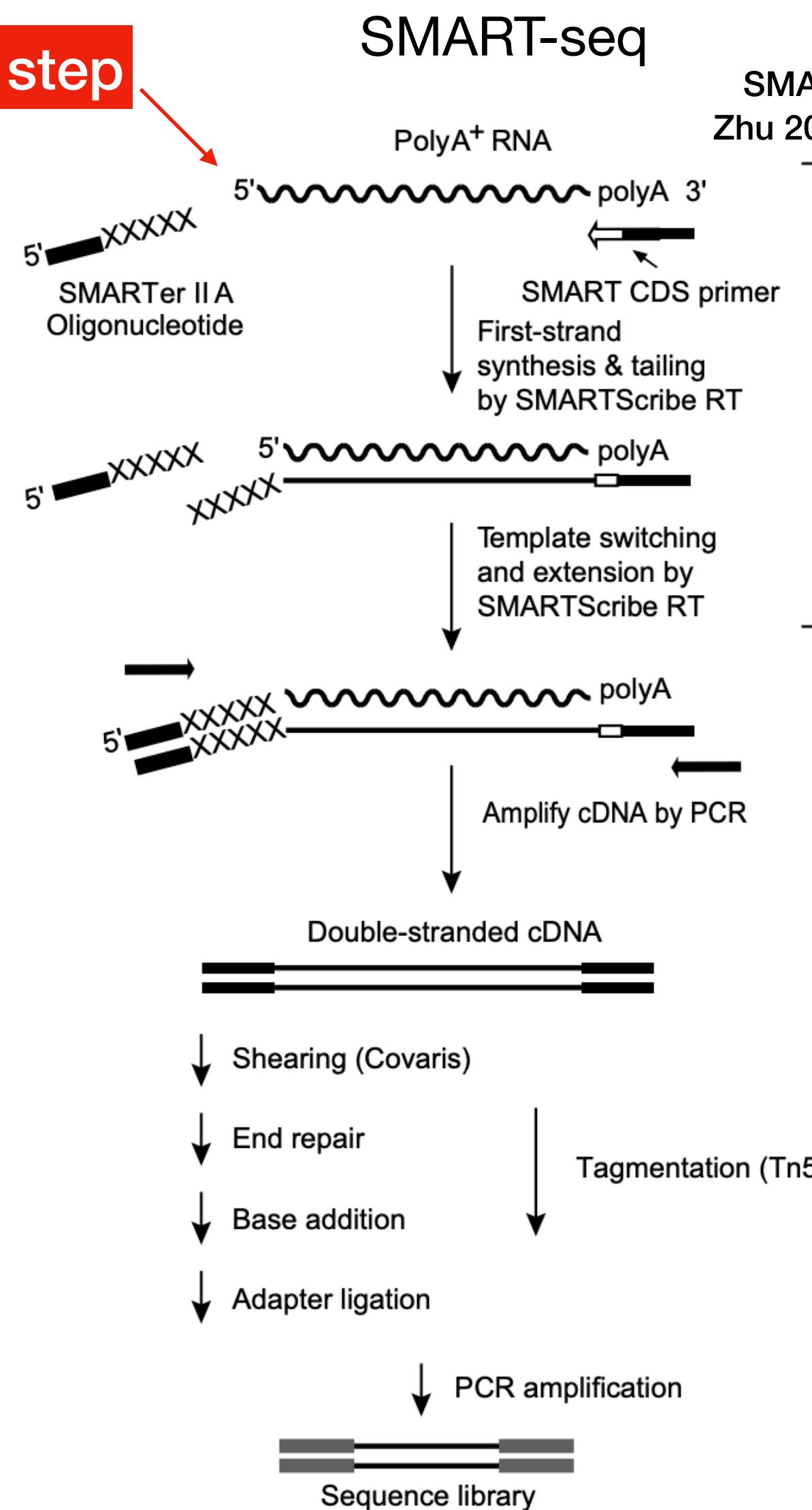


Tn5 tagmentation break DNA into small pieces, and adds barcode+adapter sequence to both ends



SMART-Seq 1, 2, 3 - Key: Full Length cDNA, Isoform (and Allele), Medium Throughput With High Coverage

Key step

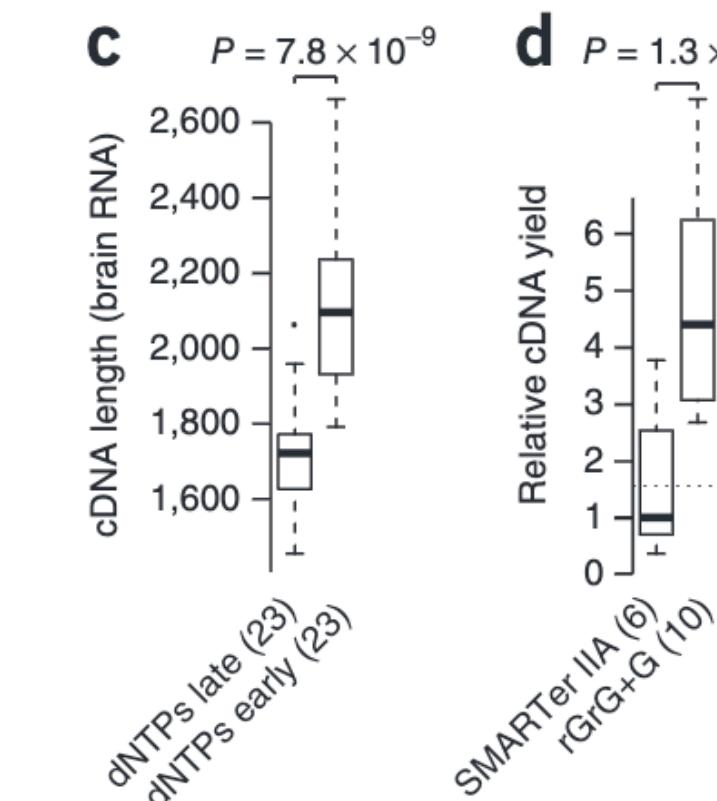
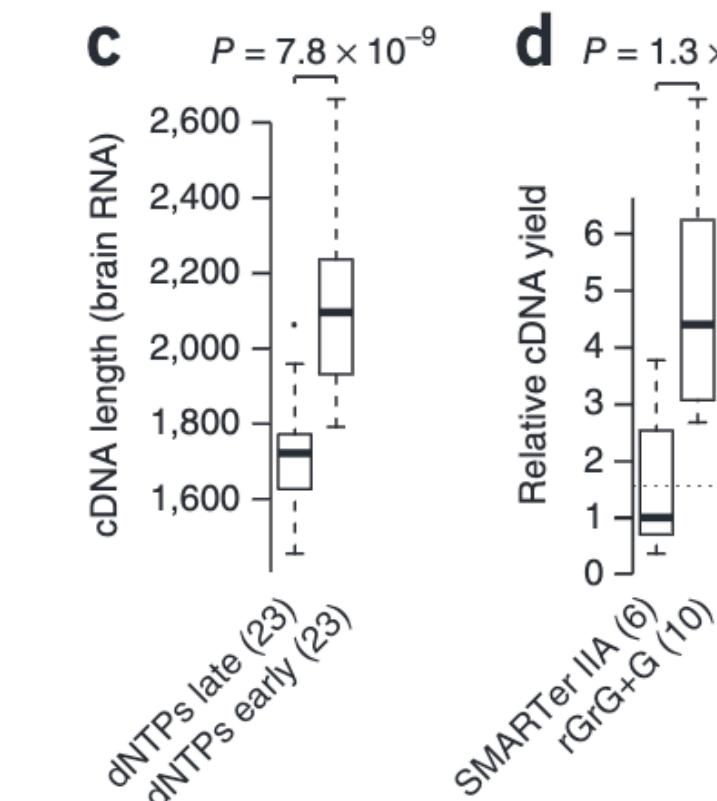
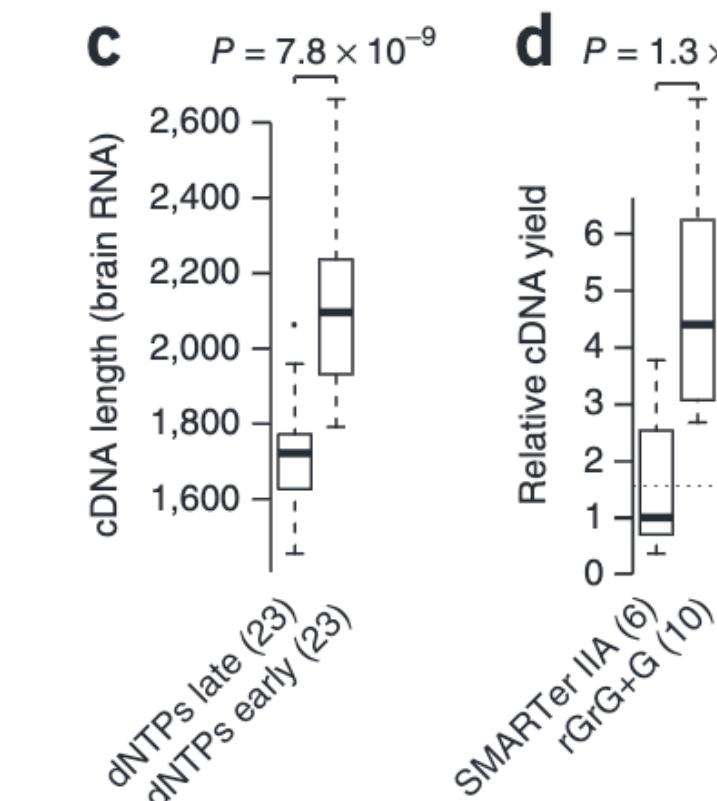
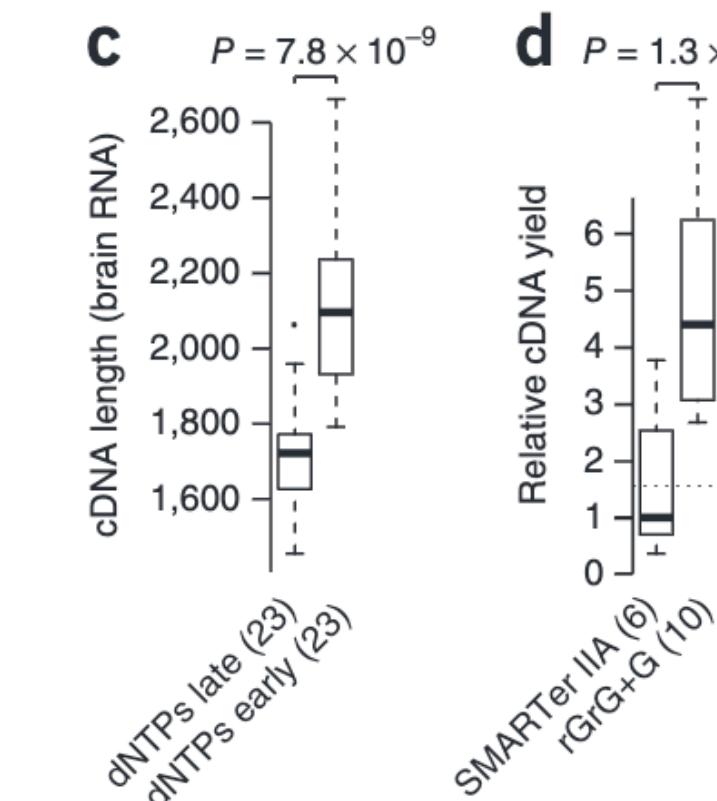
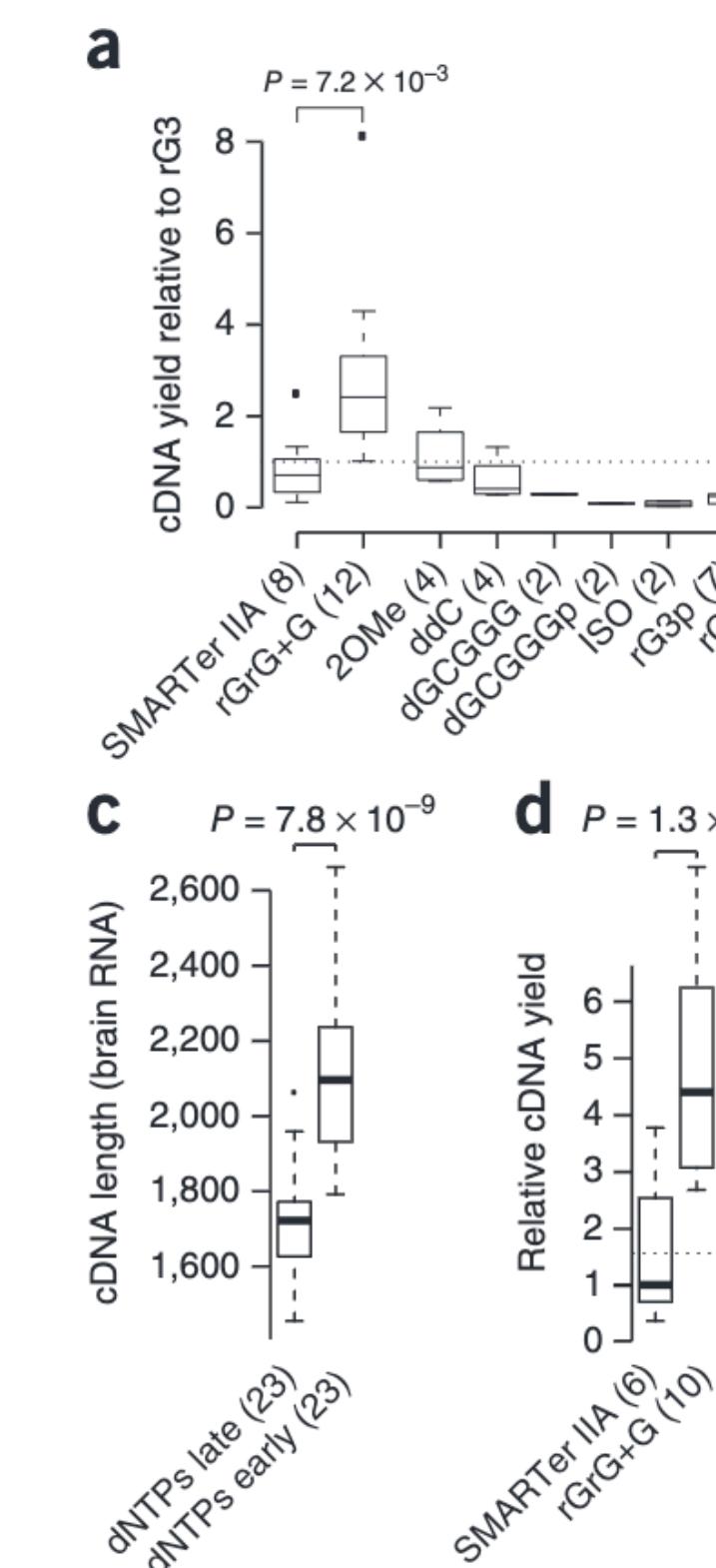


SMART-seq

SMART technology:
Zhu 2001 Biotechniques

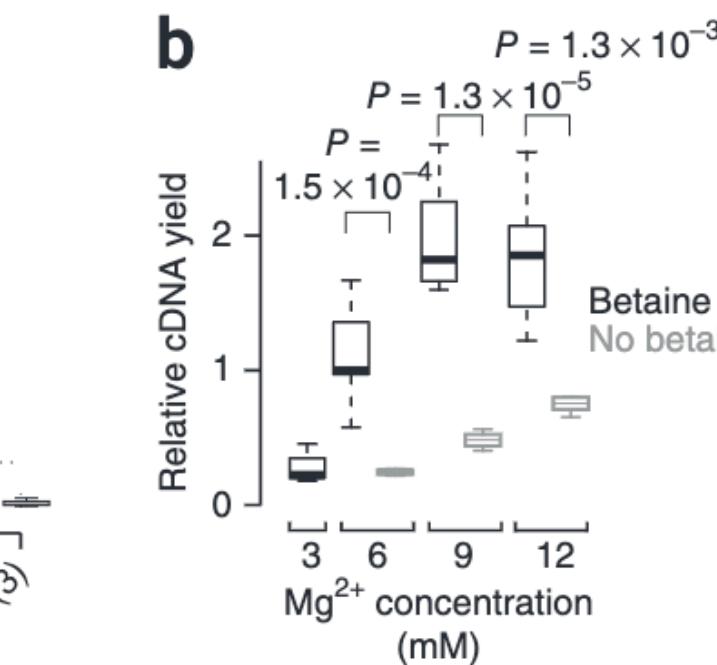
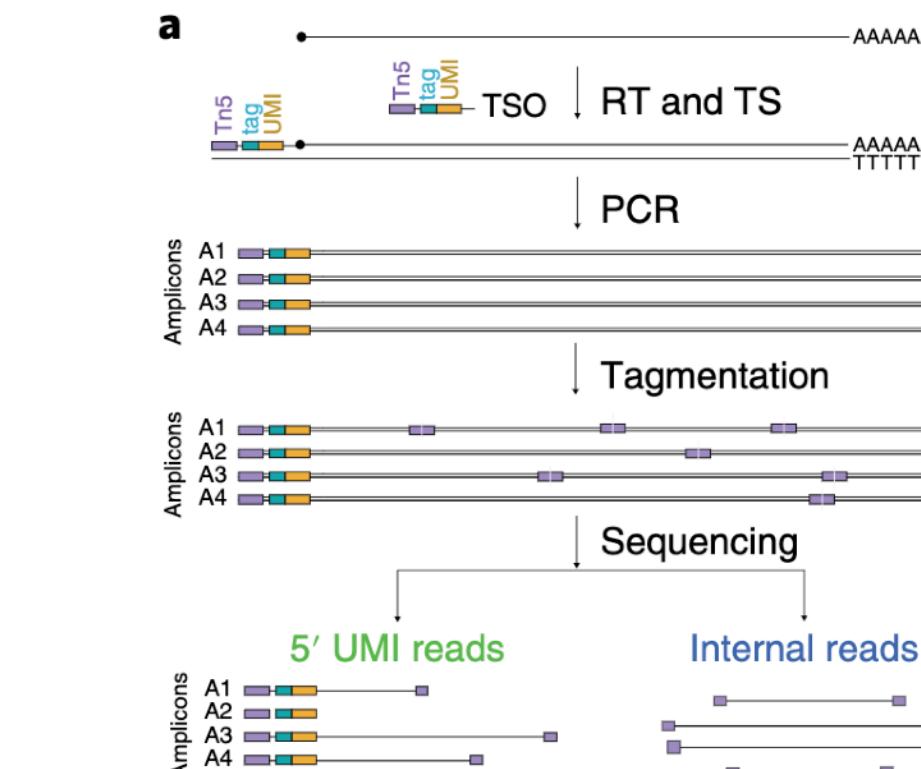
SMART-seq2

SMART-seq3



SMART-seq2 did lots of optimizations to increase sensitivity

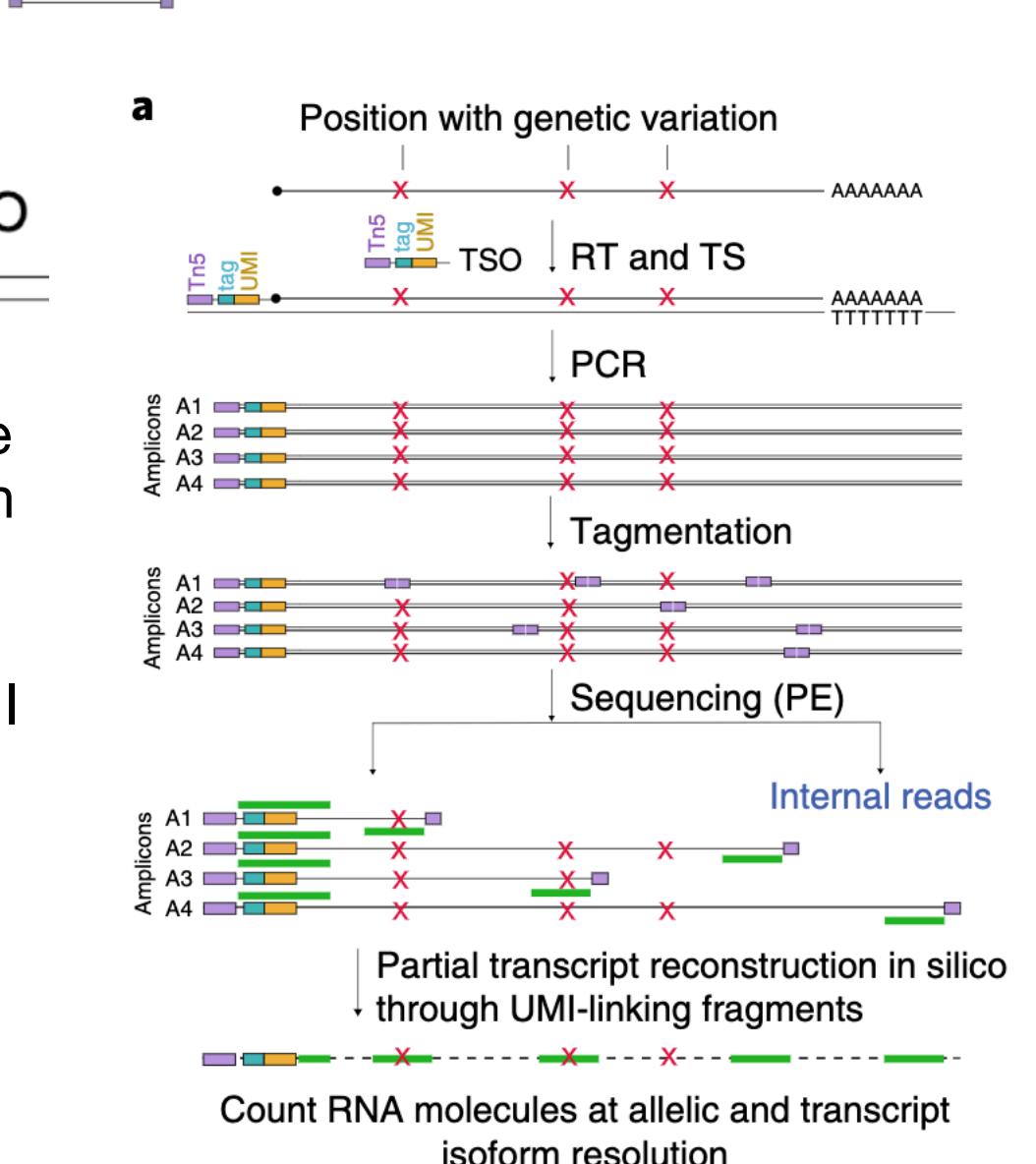
SMART-seq2



SMART-seq3 improve sensitivity, add UMI in TSO

SMART-seq3 has UMI reads and internal reads

Can look for allelic origin and transcript isoform specificity



Ramskold 2012 Nat. Biotech.

Picelli 2013 Nat. Method

Hagemann-Jensen 2020 Nat. Biotech.

Study Isoform and allele (through UMI reads)

SMART-Seq 1, 2, 3 - How a Technology Is Developed and Improved

Some thoughts

- Proof-of-concept (smart-seq): adopt an “old” technology (SMART, 2001) into a new form (single cell level).
- Detailed improvement (smart-seq2): every step, reagent, enzyme have some potential optimization, adding together, the sensitivity can be improved a lot.
- Idea multiplexing (smart-seq3): Further adopt other existing **compatible** technologies (UMI in SMART-seq3, long-read sequencing in other paper) into current technology framework.
- Future direction:
 - Multiomic: for example, SMART-seq is also compatible with ATAC and bisulfite conversion.
 - With the technology become more and more complex, more detailed improvement is needed for **specific tissue and specific question**.

A molecular technology is based on numerous building blocks.

A novel technology is usually not that “novel”.

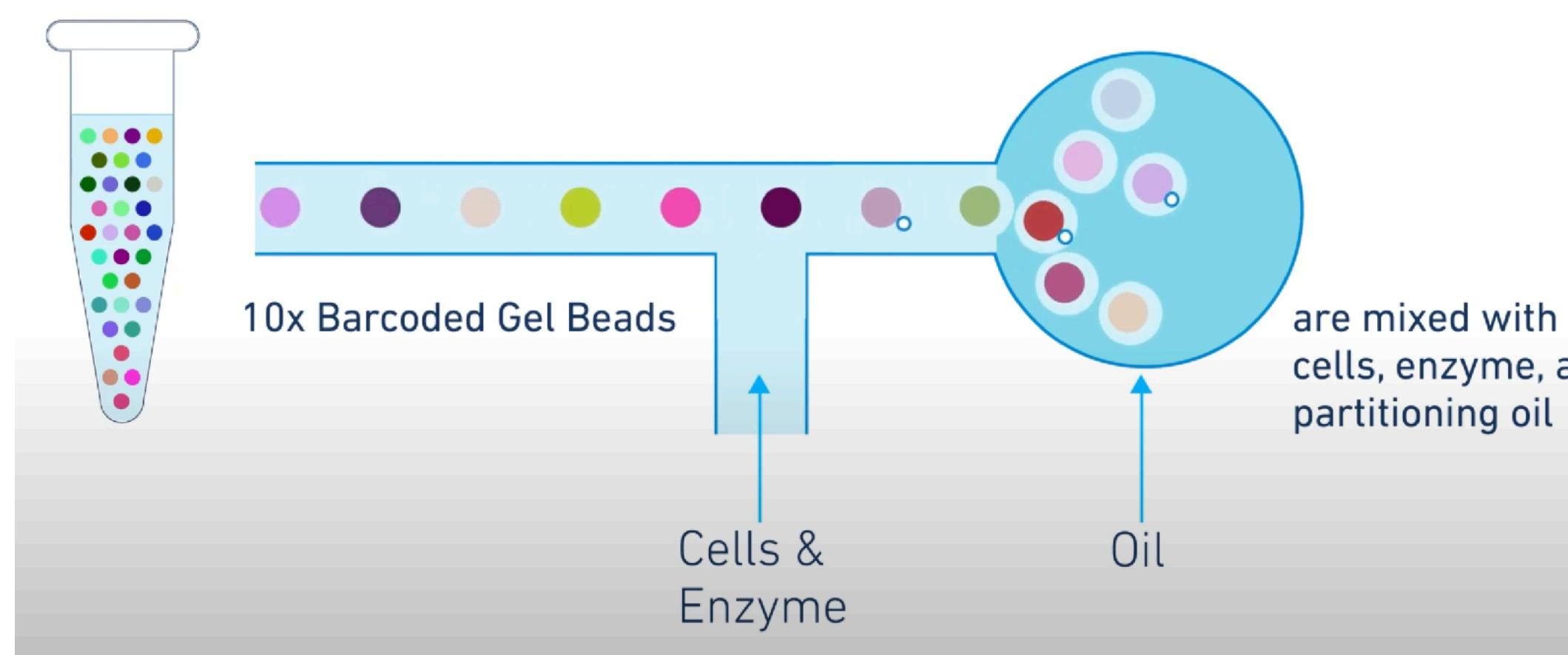
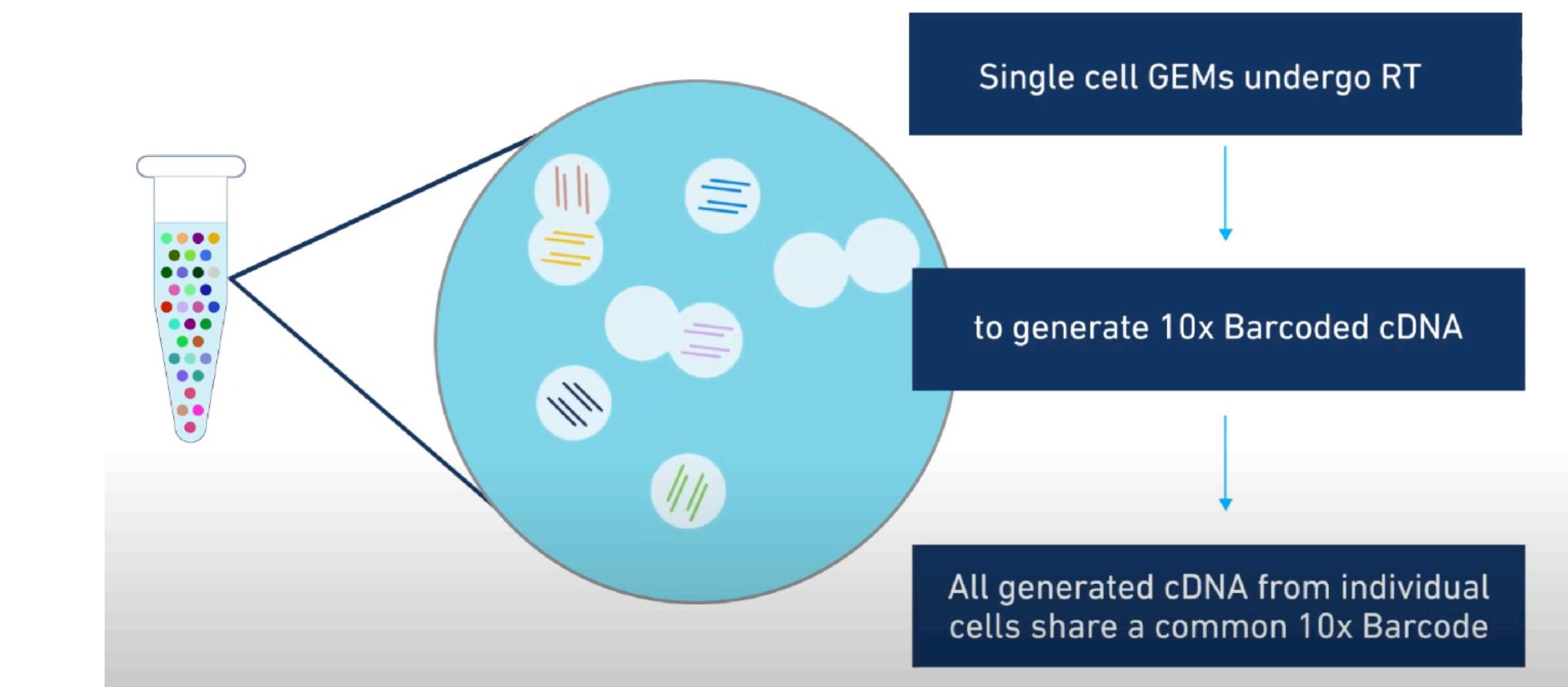
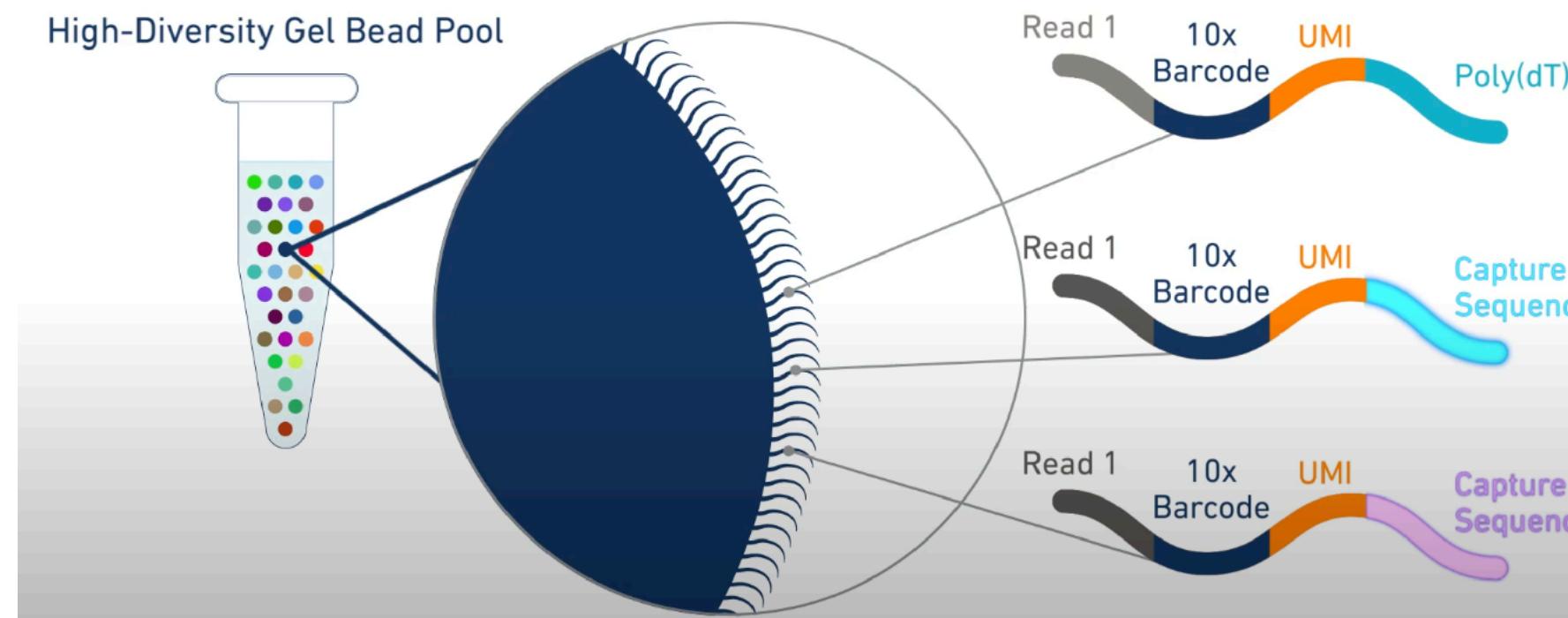
The novelty usually comes from the **new combination** of existing building blocks to achieve **new application**.

Making it work is one thing, making it work efficiently is another thing. Usually research labs put more efforts on the first one, while companies put more efforts on the second.

Question driven technology improvement is more meaningful.

10X scRNA-Seq 3' Version - 3' Enriched, UMI, Usually Hard To Tell Isoform, High Throughput With Low Coverage

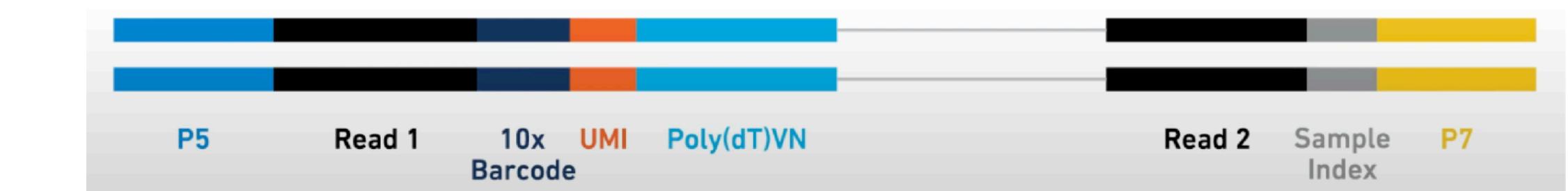
and additional feature information using capture sequences.



10X 3' version

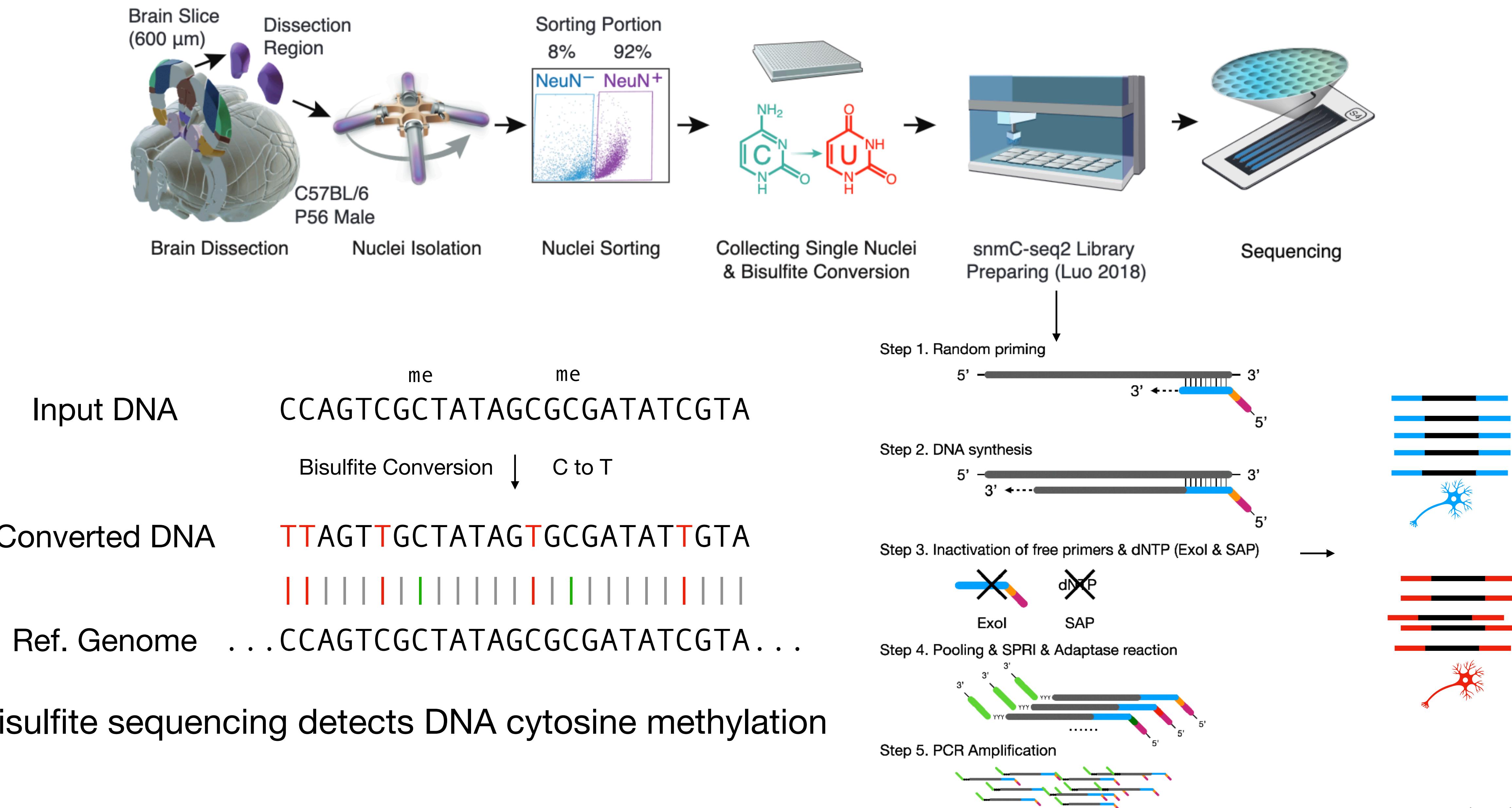
R1 contains cell barcode and UMI info for read assignment

R2 contains actual RNA sequence for mapping



Watch this video for 10X scRNA sequencing: <https://www.youtube.com/watch?v=XTZD1eNoRtc>

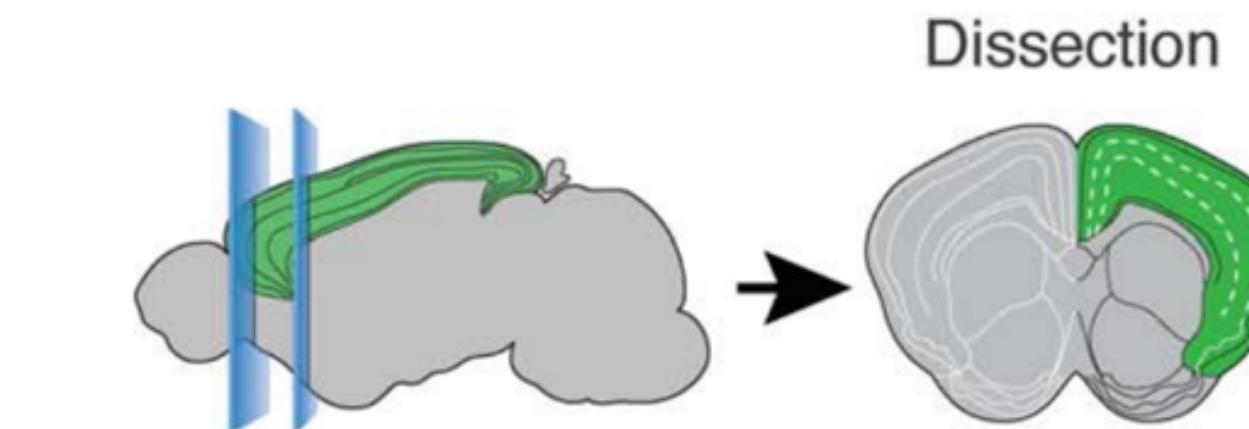
snmC-Seq2 - Whole Genome Amplification With Random Primer, Medium Throughput High Coverage



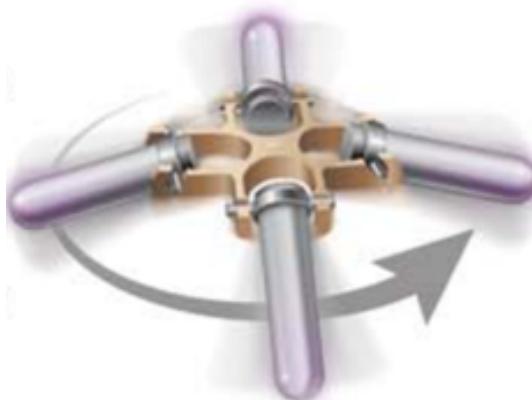
Bisulfite sequencing detects DNA cytosine methylation

snmC-Seq2 Experiments in our Lab

Dissociate single cells



Nuclei isolation



BD Influx Cell Sorter®
high-speed sorting,
maintaining cell viability
and functionality, aseptic
HEPA-filtered enclosure,
Computerized Cell
Deposition Unit (CCDU).



Salk FCC Core

M. Margarita
Behrens

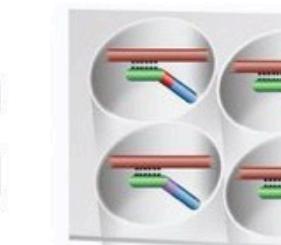
Jacinta
Lucero

Dissection

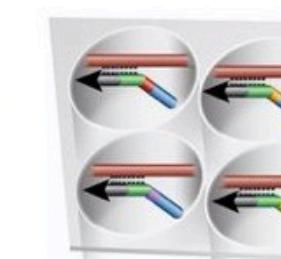
Automated Library Preparation



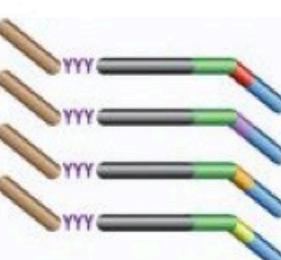
Random priming



cDNA synthesis



Barcoding & Pooling



Amplification

Thermo Fisher ProFlex
PCR System®
2x384-well plates



Ecker lab (Anna B.)

High-throughput massively parallel sequencing

Pilot QC run
on miseq



Ultra-deep sequencing on
Illumina Novaseq 6000
platform; S4 flowcell



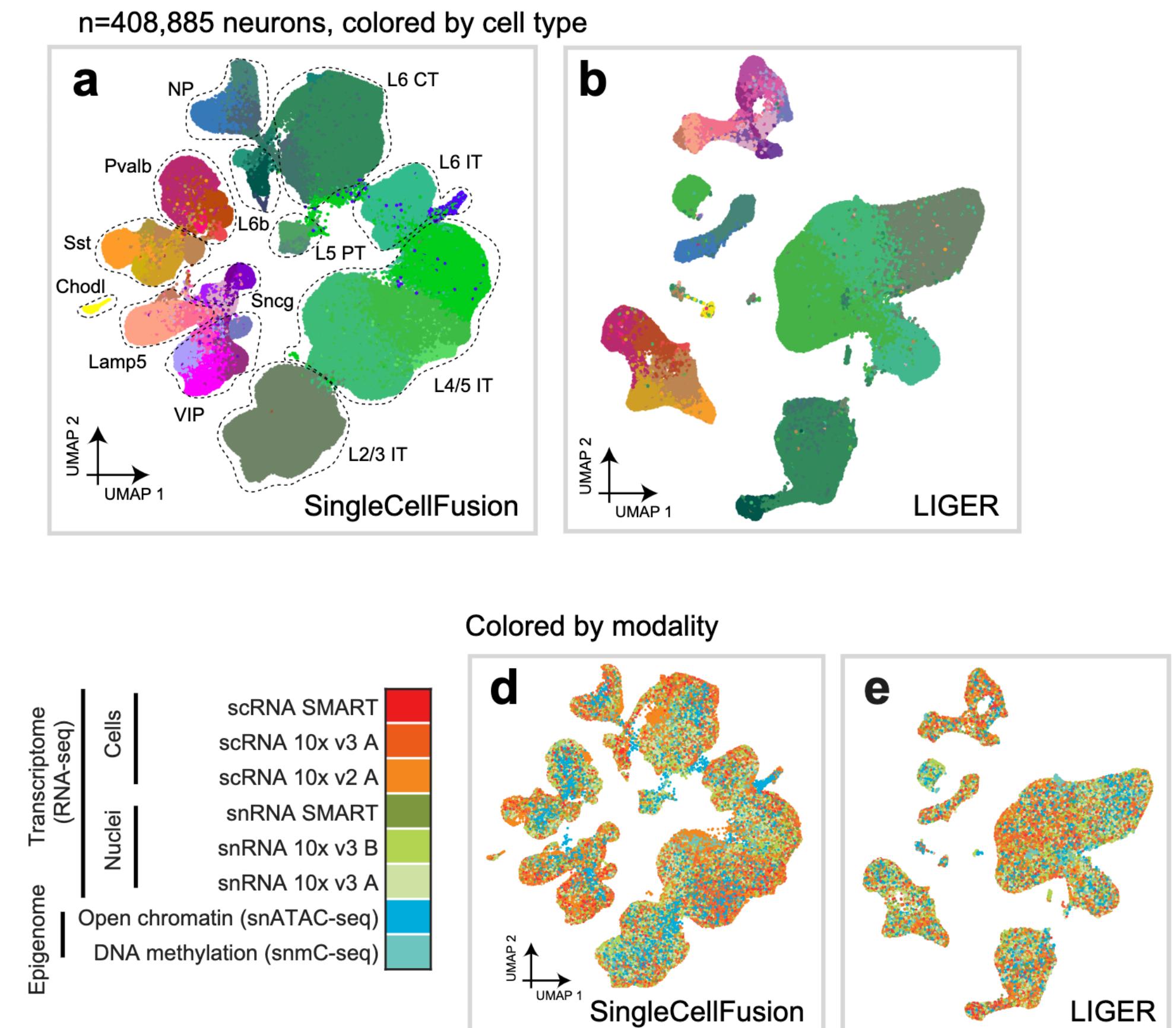
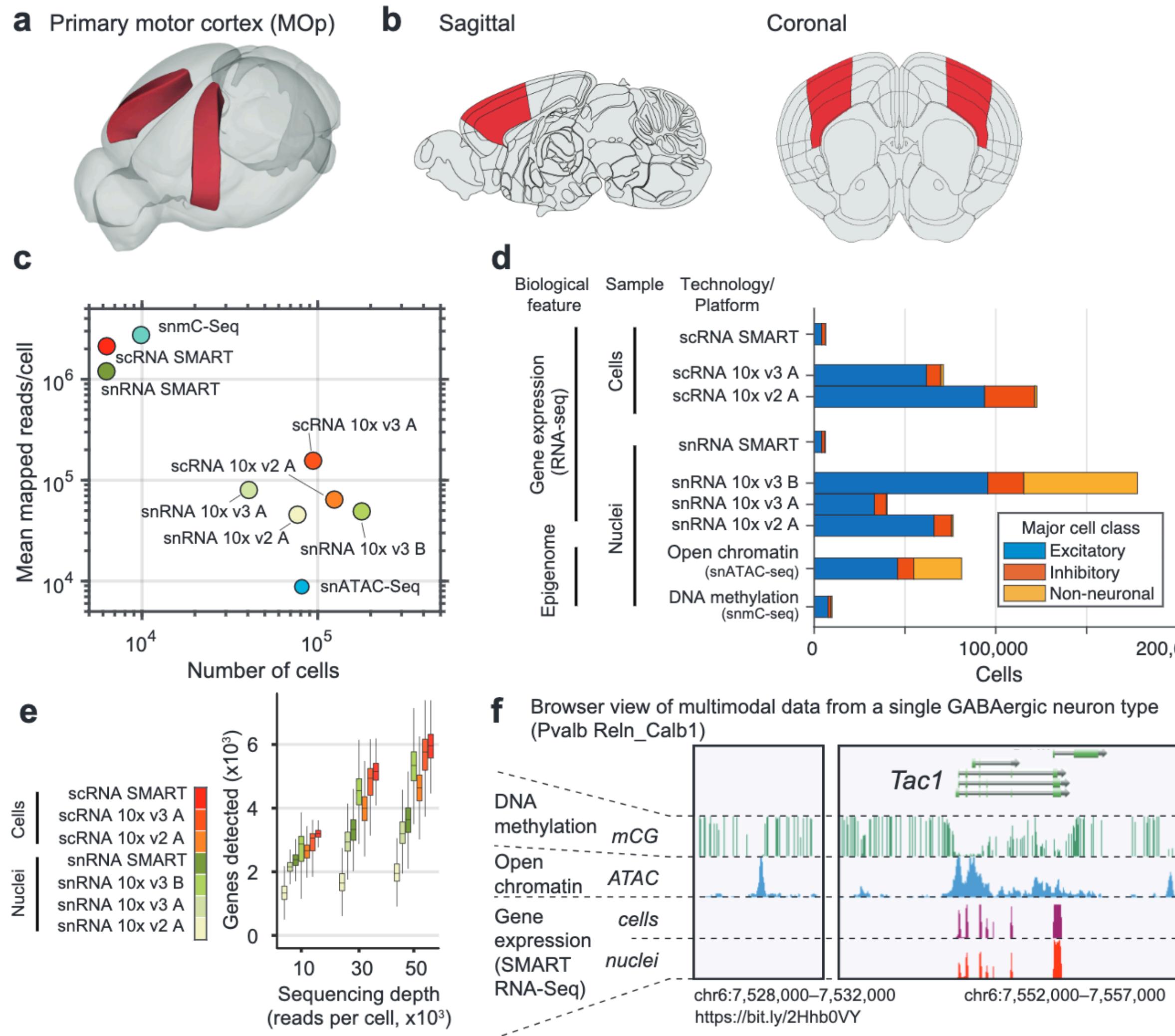
Joe Nery

Illumina Novaseq 6000®:
Single run with
16 x 384-well plates
(6,144 cells) per run (48
hours); ~3.5 terabases
from 2x150 (PE) reads

Ecker lab (Joe, N.)

**Example: compare and integrate
multiple different technologies on
the same tissue**

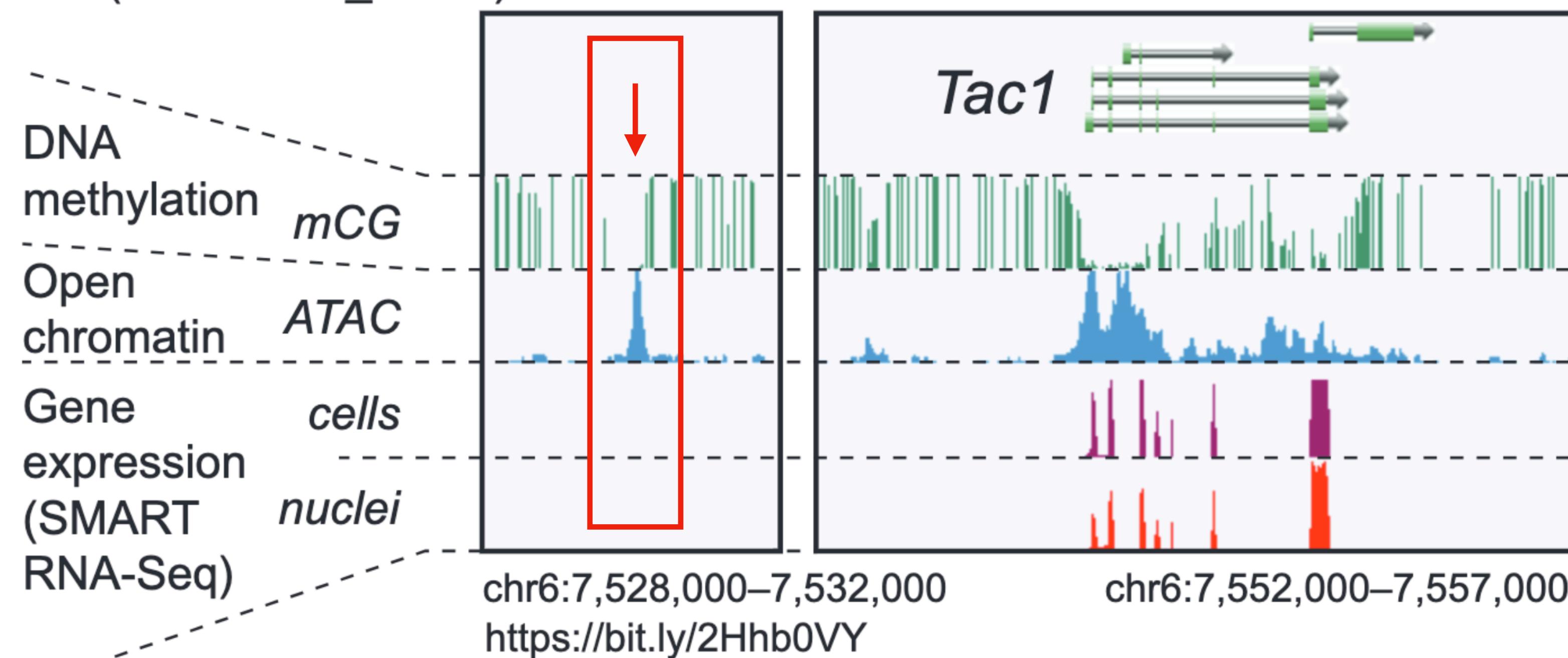
Molecular Taxonomy of the Mouse Brain Primary Motor Cortex



Integration of transcriptomic and epigenomic (DNA methylation and accessible chromatin) profiling identified 56 cell types in mouse primary motor cortex.

Epigenome Profiling Predicts Cis-Regulatory Elements

f Browser view of multimodal data from a single GABAergic neuron type
(Pvalb Reln_Calb1)



Epigenome Profiling

Transcriptome Profiling

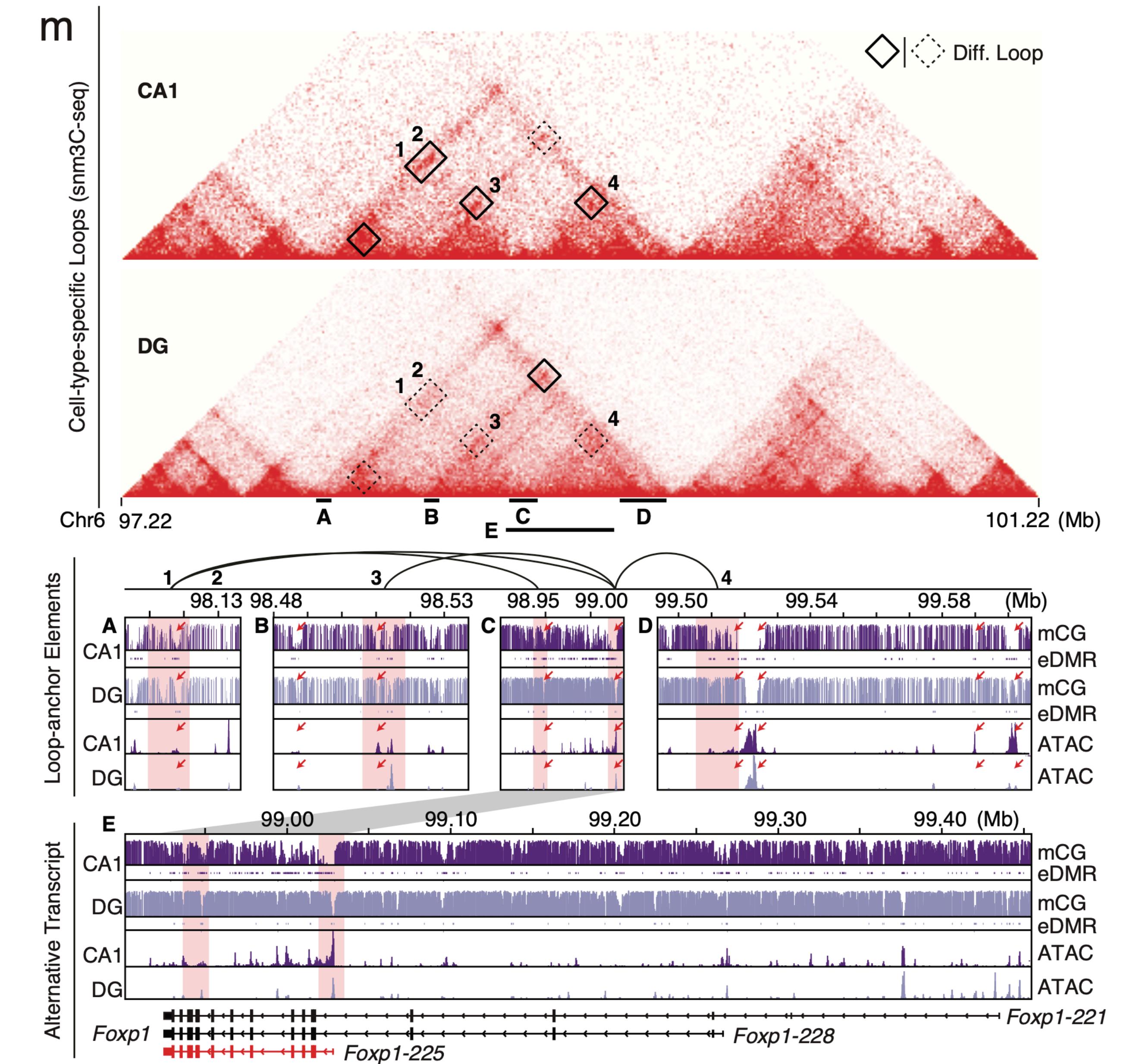
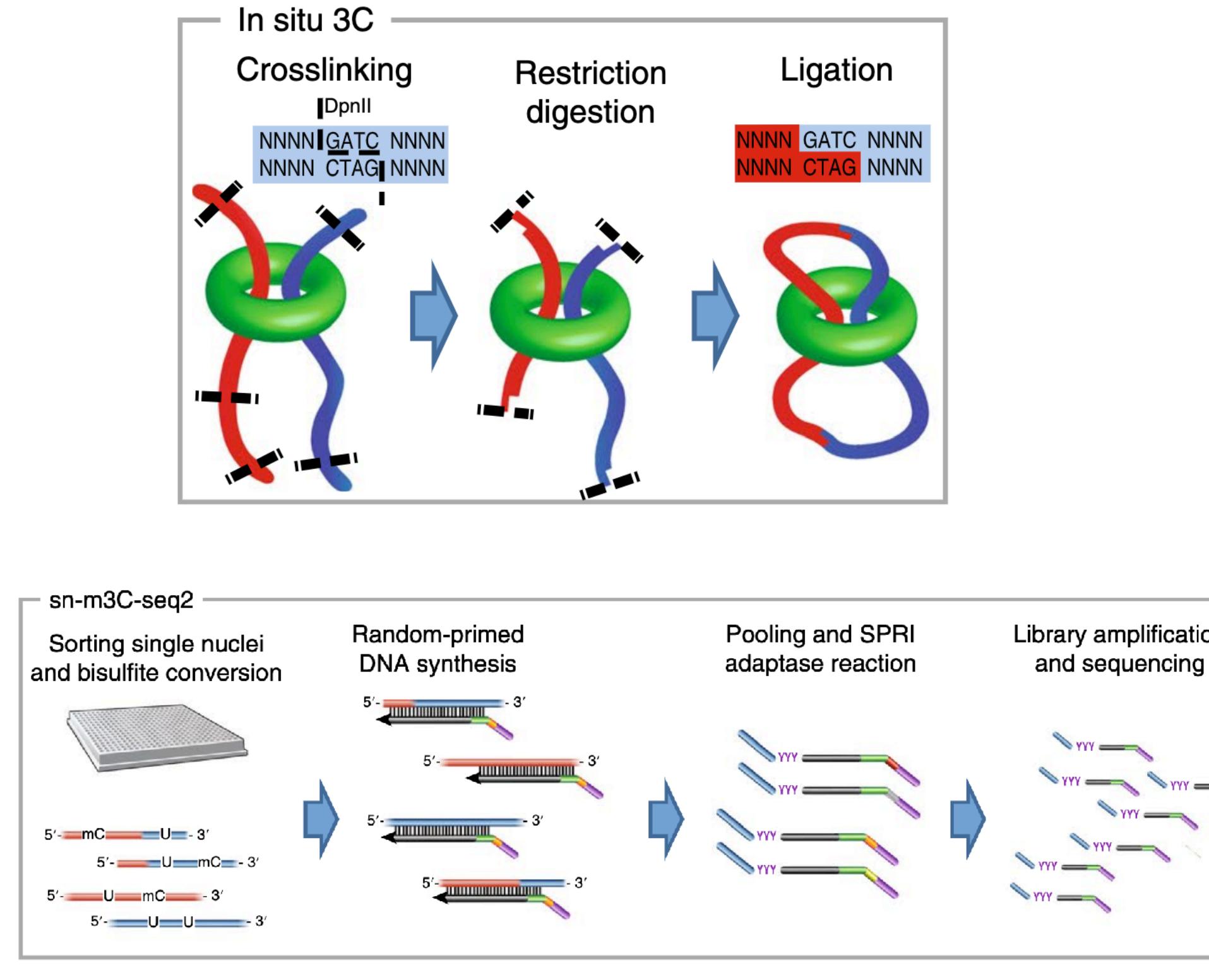
Gene & Cis-regulatory elements

Gene

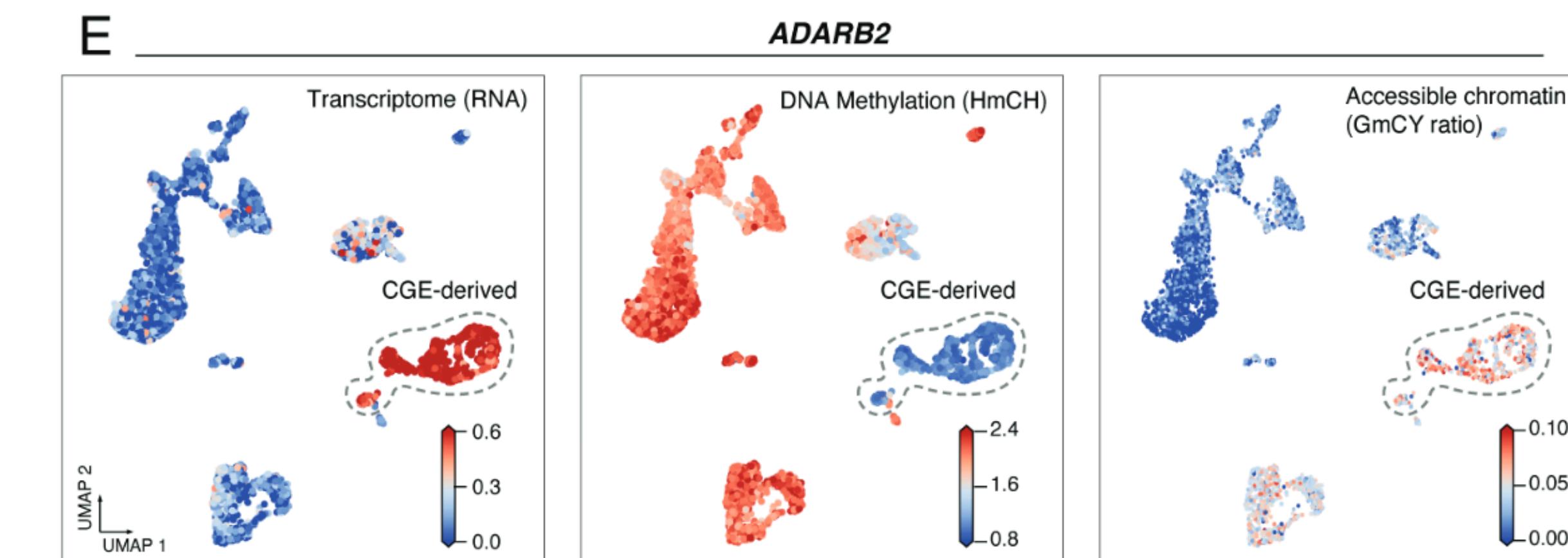
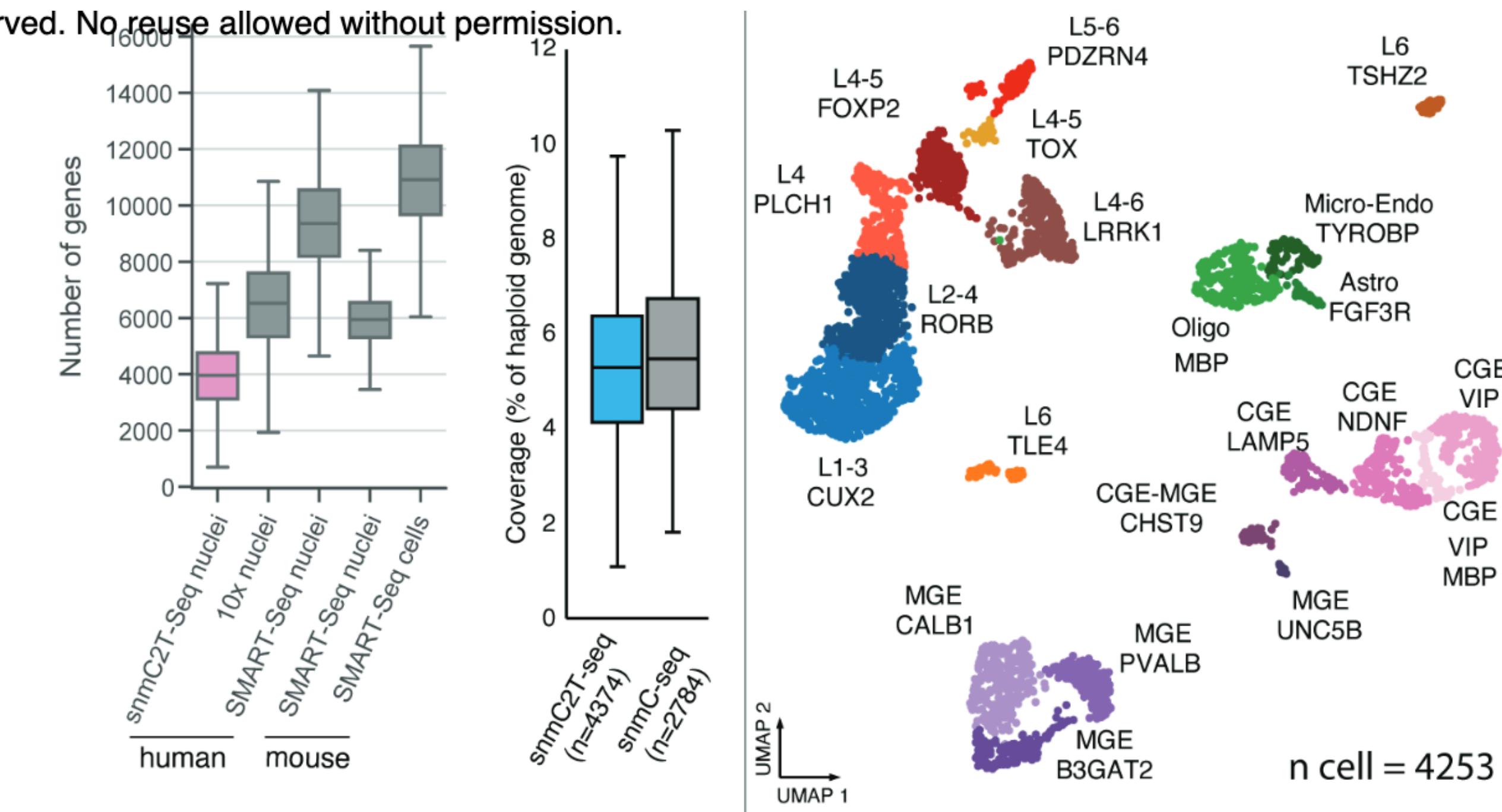
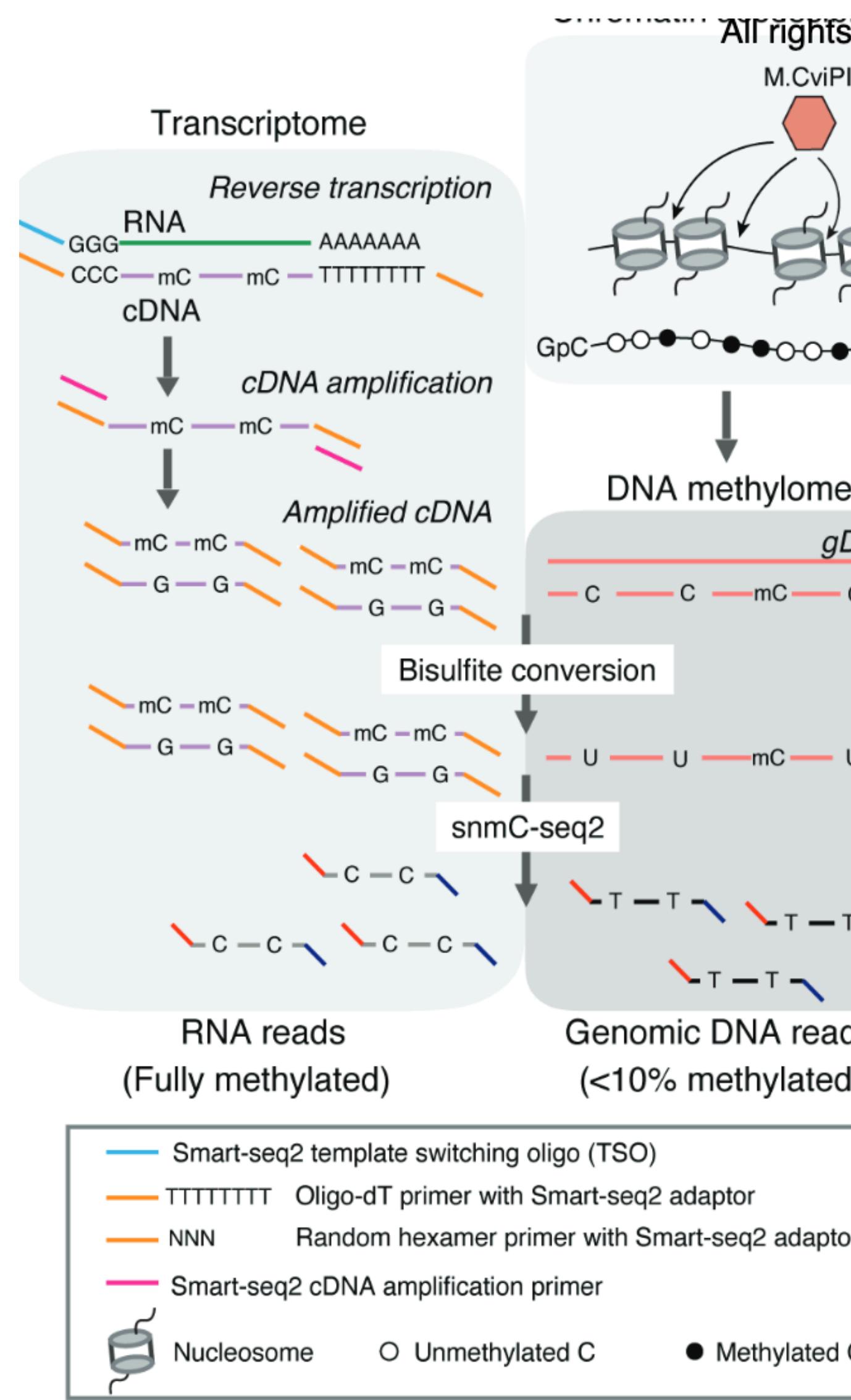
Future Directions (Tech. Dev. Aspect)

- Multi-omic technology
- Spatial omics

snm3C-Seq: Chromatin Conformation + DNA Methylome



snmC2T-Seq: DNA Methylome + Transcriptome + Chromatin Accessibility



Future Directions (Tech. Dev. Aspect)

- Multi-omic technology
- Spatial omics

Spatial Omics Summary

- Modality been achieved:
 - Mainly transcriptome,
 - Multiplexed protein level detection via barcoded antibody, but not reaching omics level.
- Two main tech. genres of spatial (transcript)omic tech.
- **Image/FISH based** (MERFISH, SeqFISH+)
- **Sequencing based:**
 - **In-situ FISH-like sequencing:** FISSEQ, STARMAP
 - **In-situ spatial barcoding + NGS sequencing:** Slide-Seq, 10X Visium, DBiT-seq
- Some thoughts:
 - Still in its early phase, but developing super fast.
 - Fancy tech. require specialized tissue preparing and equipment handling knowledge, cost is still high and time consuming. Most papers still come from original tech. dev. lab.
 - Commercialized tech. haven't really reach single cell level, or haven't really reach omic level
 - Other omic (e.g. epigenomic) is still under development

- Spatial omic summary resource:

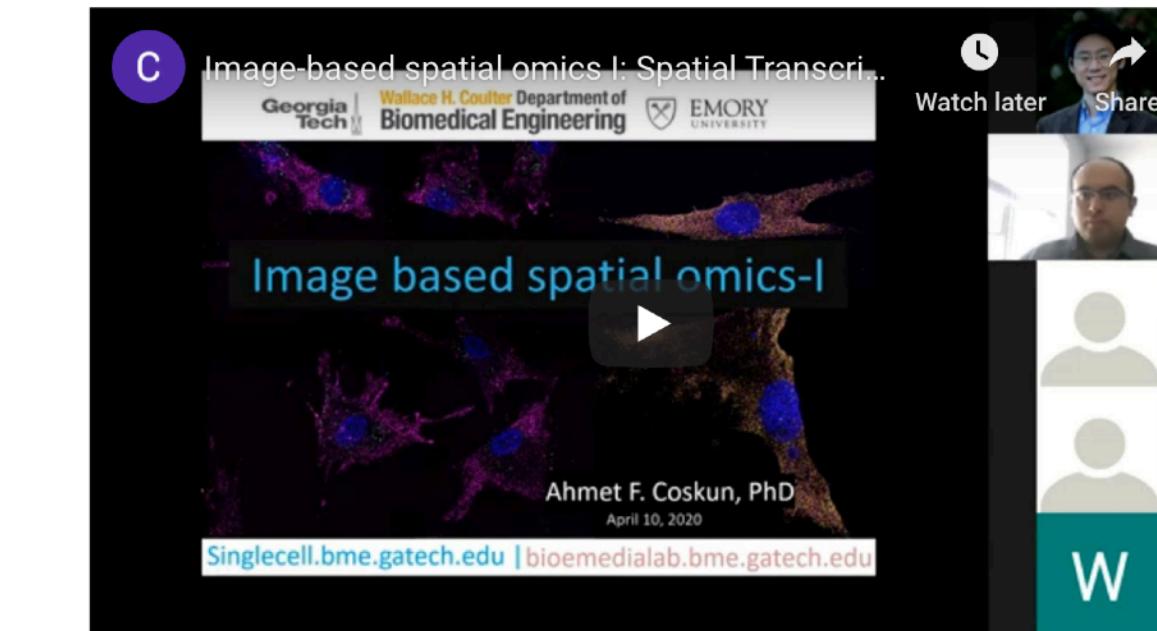
<http://spatialomics.net/>

<https://spacetx-starfish.readthedocs.io/en/latest/>

- Highly recommend talks for summary of current status:



<http://spatialomics.net/talks/>



“Progress in science depends on new techniques, new discoveries and new ideas, probably in that order.”

–**Sydney Brenner**