

# Introduction to Single Cell RNA-sequencing: a practical guideline

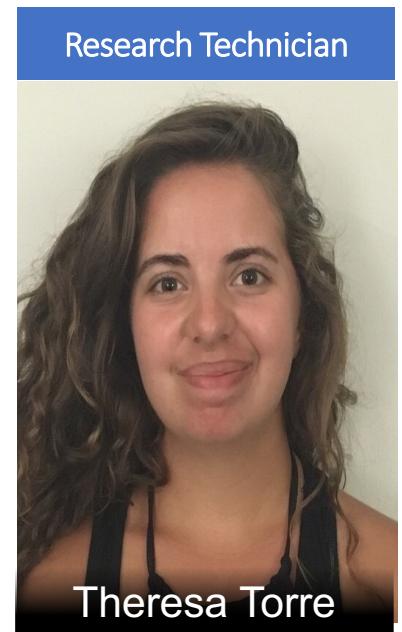
Mandovi Chatterjee, Ph.D.

Director, Single Cell Core at Harvard Medical School

Research Associate, Dept. of Systems Biology

# Single Cell Core @ HMS

**Mission: Enable breakthrough discoveries by assisting in the design, execution, and interpretation of single cell genomics assays, utilizing cutting-edge, state-of-the-art tools**



## Faculty Advisors:

Allon M. Klein, PhD

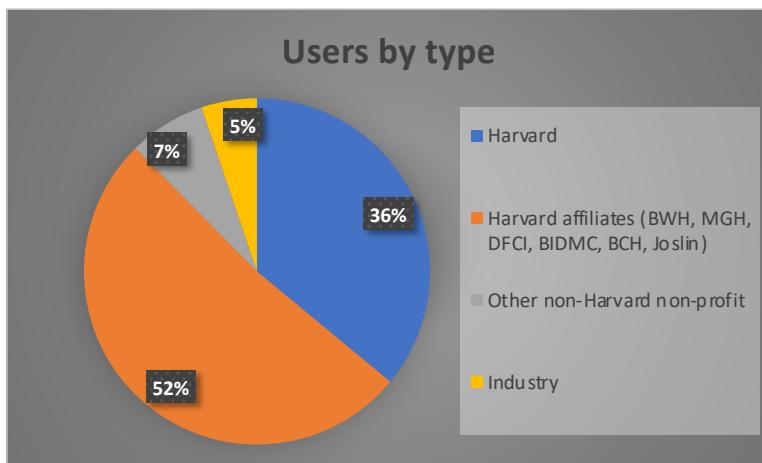
Jeffrey Moffitt, PhD

Christophe Benoist, MD, PhD

Harvard Medical School  
Armenise Building #517, Boston, MA

<https://singlecellcore.hms.harvard.edu>

# Single Cell Core Facility



Calendar Year	2022	2023
Number of new projects	30	29
Number of users	39	44
Number of labs	31	38
Number of runs	104	117
Number of cells barcoded	More than 3 million	More than 4 million
Number of samples handled/ libraries prepared	More than 350	More than 500
Number of experiments	42	48

## Key Services

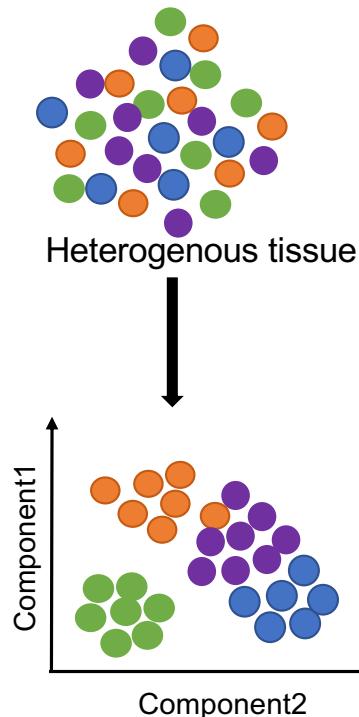
- Consultation
- Single Cell mRNA barcoding
  - 10x Chromium
  - Parse Biosciences (SPLiT-seq)
  - Fluent Biosciences (PIP-seq)
- Applications
  - scRNA-seq (Both 3' and 5')
  - scATAC-seq
  - Multiome
  - CITE-seq, hashtagging, MULTI-seq, CellPlex
- Library Preparation (both short-read and long-read)
- Single nuclei preparation
- Spatial transcriptomics (Visium and MERSCOPE) and more...

# Outline

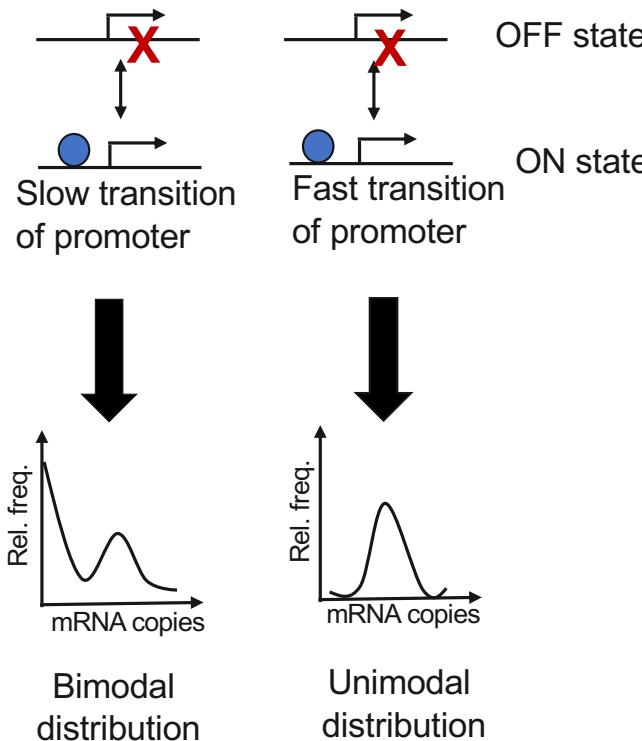
- Common applications of scRNA-seq
- Different platforms available for high-throughput scRNA-seq (in our core and some more)
- Steps in single cell RNA-seq experiments
  - Experimental design / Sample prep
    - Sample quality
    - Dissociation
    - Enrichment
    - Single nuclei vs single cell prep
    - Cryopreservation vs fixed samples
  - Encapsulation and seq depth
- Challenges- replicates, batch effect- sample pooling (hashtagging and MULTI-seq)

# Application of scRNA-seq

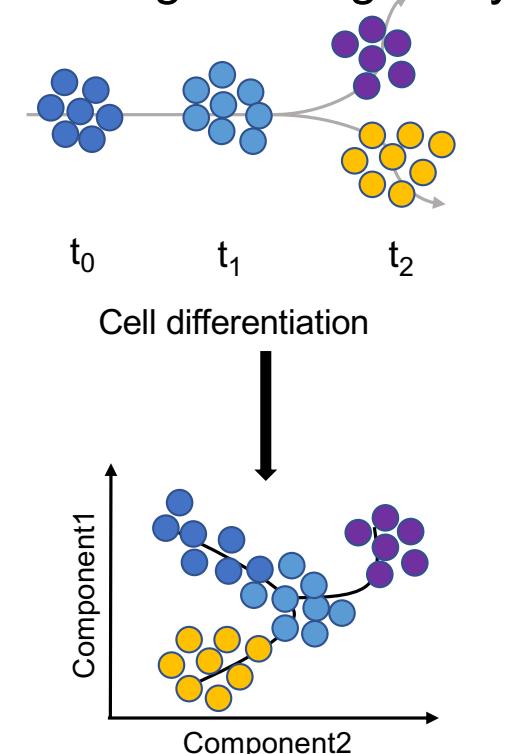
## Studying heterogeneity



## Stochastic gene expression



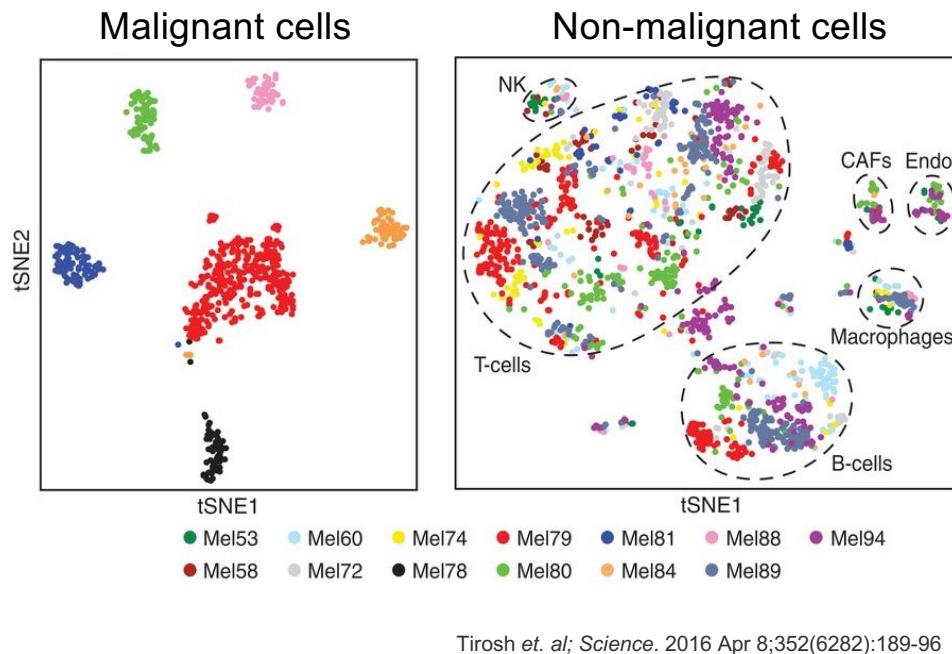
## Lineage tracing study



Adapted from:

Liu S and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Research* 2016, 5:182 (doi: 10.12688/f1000research.7223.1)  
Junker and van Oudenaarden; Every Cell Is Special: Genome-wide Studies Add a New Dimension to Single-Cell Biology; *Cell* Volume 157, Issue 1, p8–11, 27 March 2014

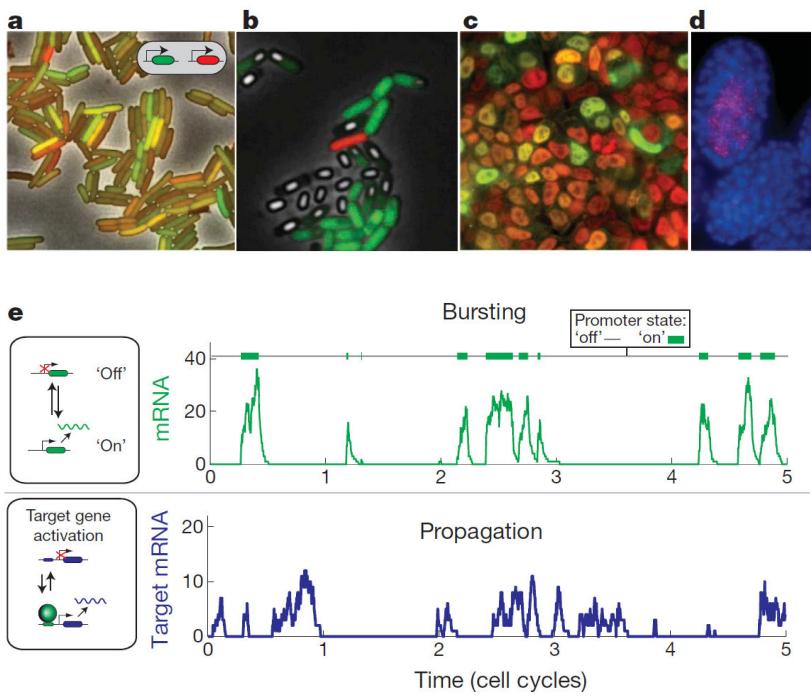
# Application of scRNA-seq: cellular heterogeneity



- Here cells were isolated from 6 melanoma patients and both malignant and non-malignant cells (both stromal and immune cells) were profiled.
- Malignant cells from 6 different tumors formed separate clusters on a t-SNE plot, suggesting high degree of inter-tumor heterogeneity.
- Non-malignant cells clustered based on cell types. The clusters were annotated as T cells, B cells, NK cells, macrophages, endothelial cells and cancer associated fibroblasts (CAFs).

# Application of scRNA-seq: Stochastic gene expression

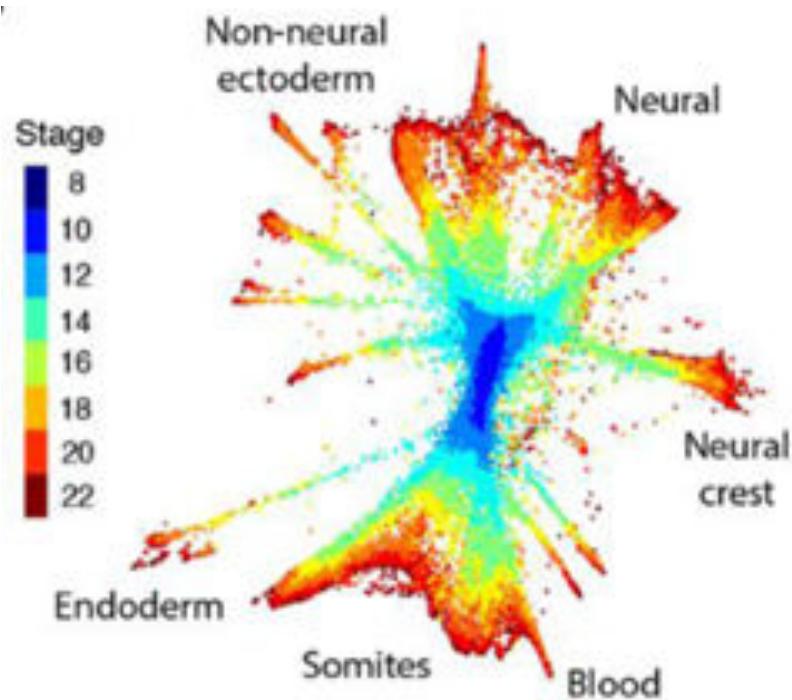
## Stochastic gene expression study



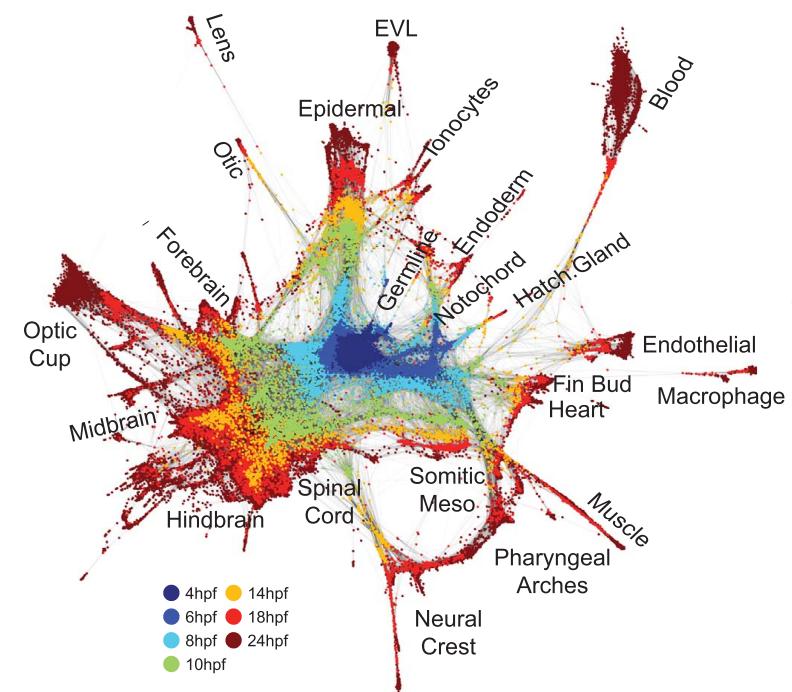
- Genes fluctuate between “On” and “Off” promoter state. Therefore, their expressions are heterogeneous and “bursty”.
- Regulated by a complex interaction between different transcription factors and chromatin-modifying proteins, which are often transient.
- Stochasticity in one gene can propagate to generate more stochasticity in downstream genes.
- In-depth analysis of scRNA-seq data can reveal such complexity.

# Application of scRNA-seq: Lineage tracing

Frog

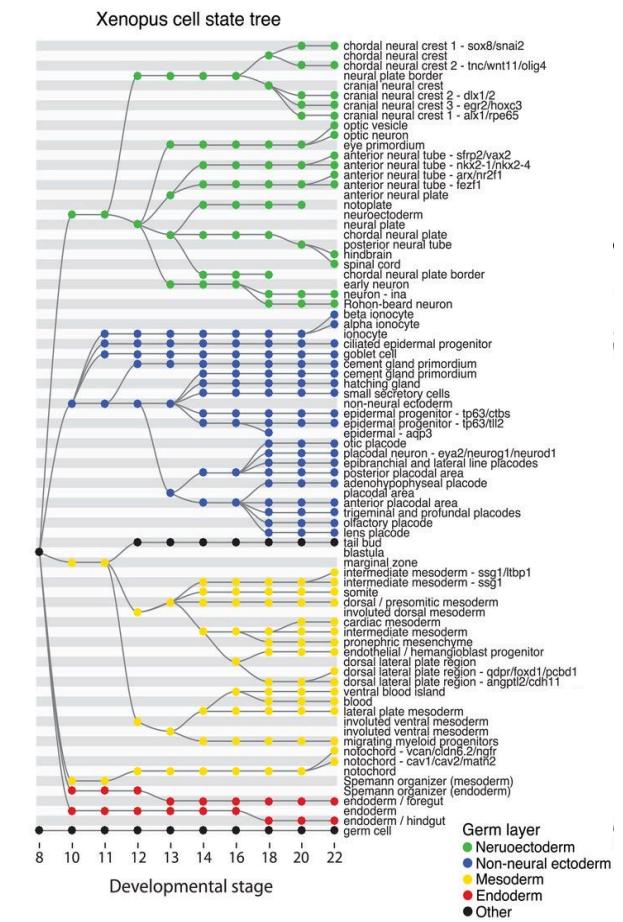
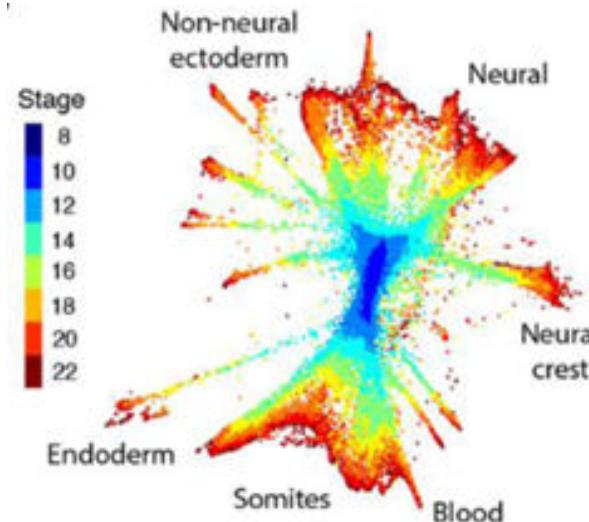
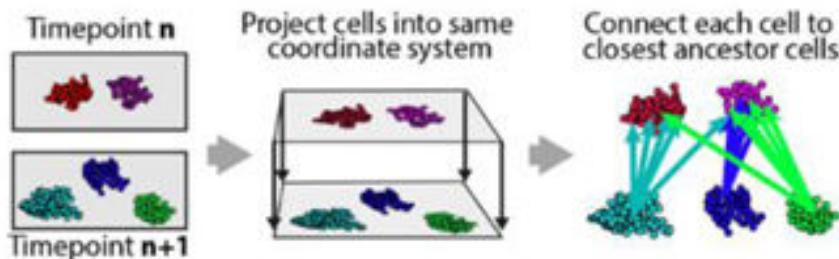


Zebrafish

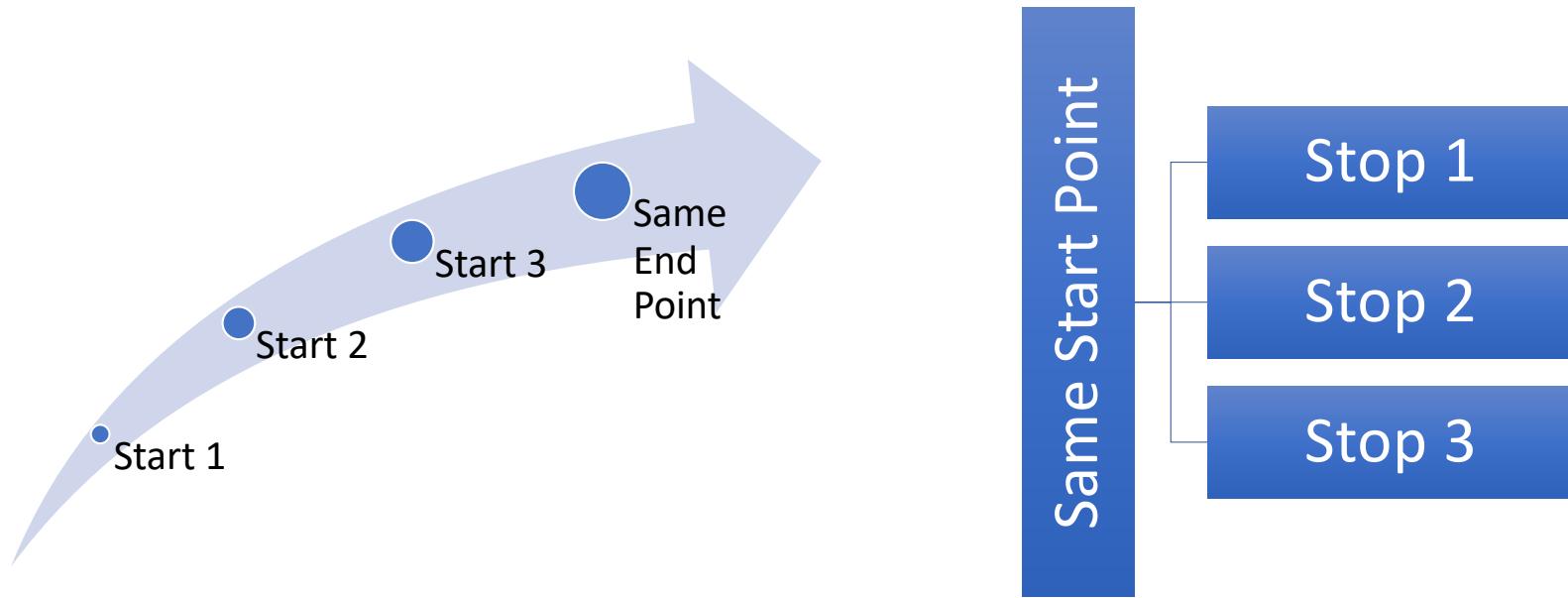


JA. Briggs et al. Science 01 Jun 2018  
DE Wagner et al. Science 01 Jun 2018

# Application of scRNA-seq: Lineage tracing

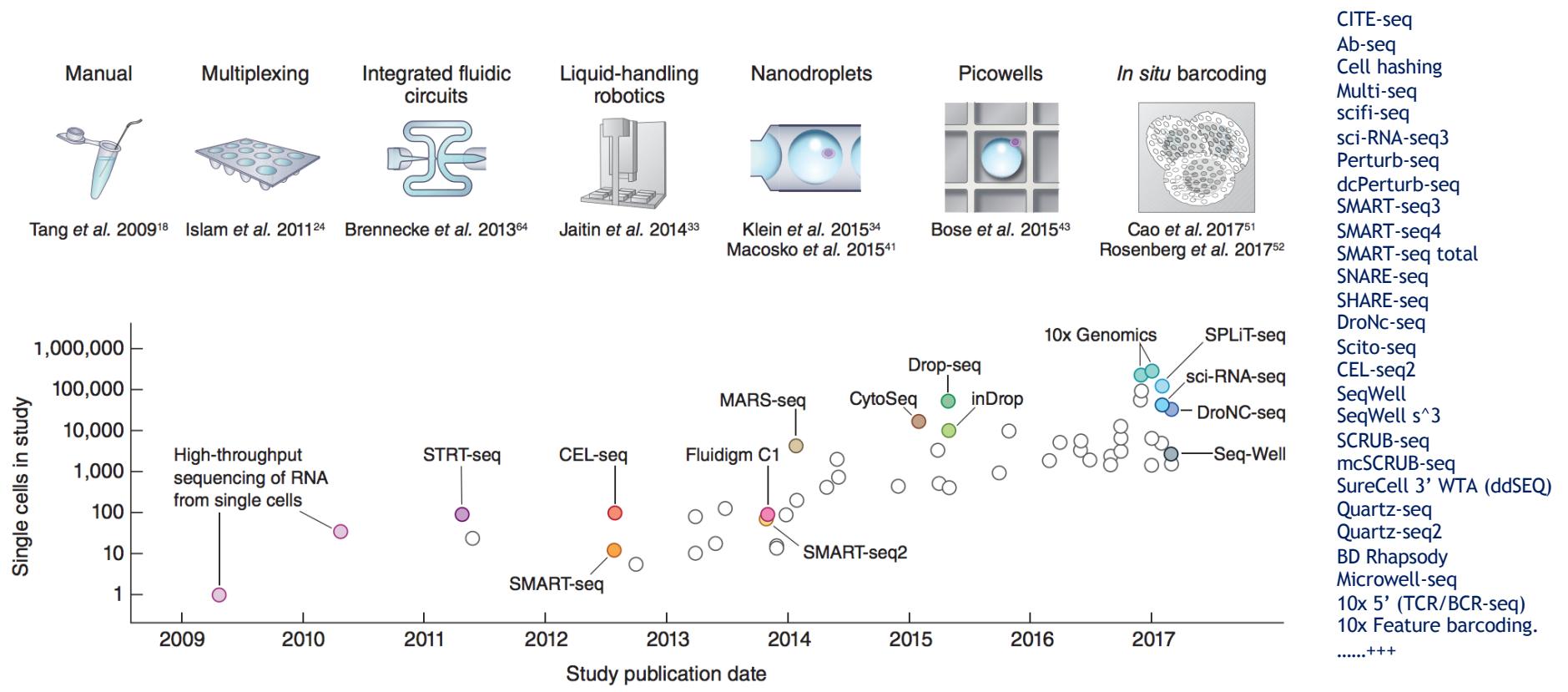


## Considerations: Time course or Development Experiments



- Collect all samples and prep libraries together in one batch.
- Biological duplicates (at minimum)

# Exponential increase in throughput and new smart tech



Svennson et al., *Nature Protocols* (2018)

# Different platforms available for scRNA-seq

## In wells

(FACS-assisted)

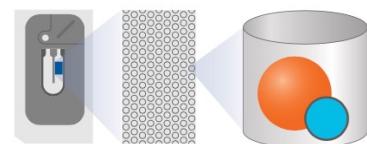


CEL-Seq  
MARS-Seq  
SMART-Seq  
SCRB-Seq  
Quartz-Seq

## In nano/microwells

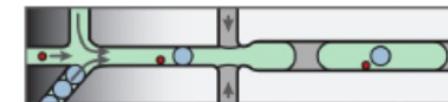
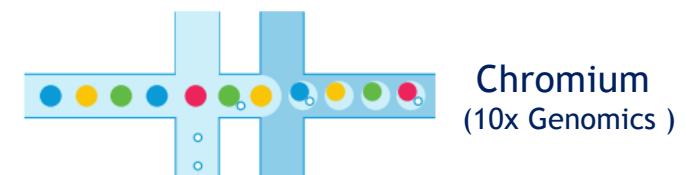


Seq-Well

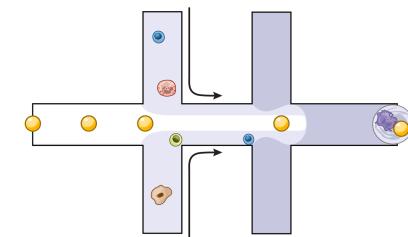


BD Rhapsody

## In droplets

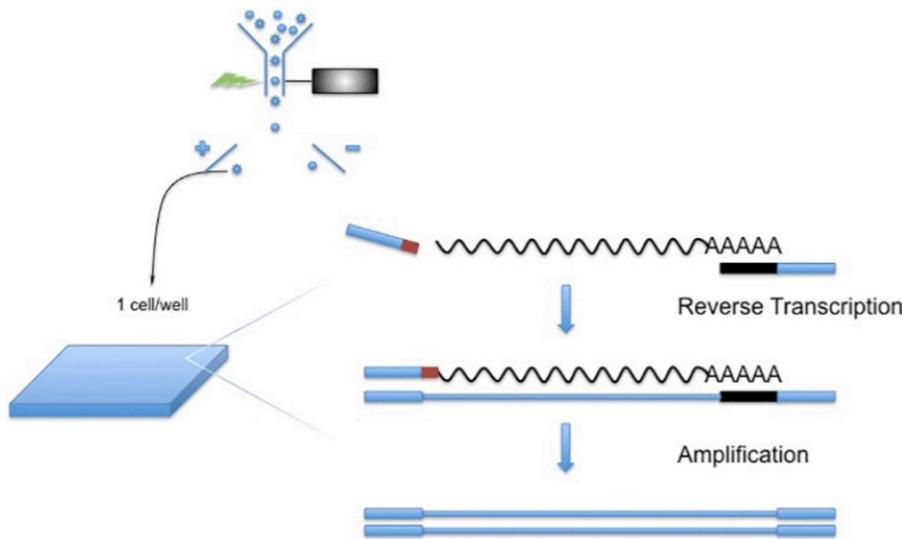


InDrops



DropSeq

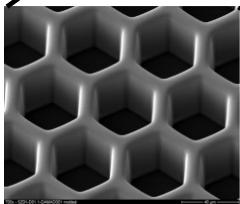
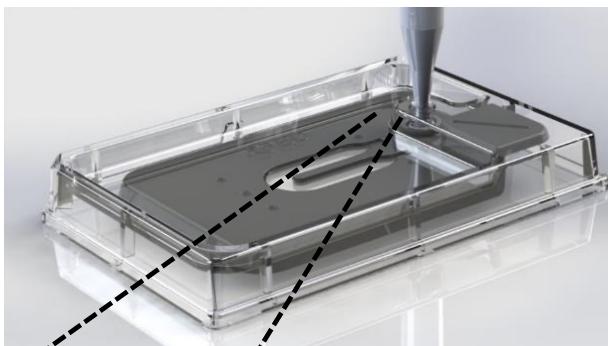
# SMART-seq (v3): Full length transcript



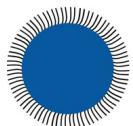
- Sort cells of interest into single well.
- Only single cell method that gives full transcript information.
- Currently best option for low cell number samples. (100's – 1,000's)

# Microwell-based platform: BD Rhapsody

Cartridge



- Single use
- Easy to load
- 200k+ microwells



- Magnetic bead with immobilized oligos

Scanner



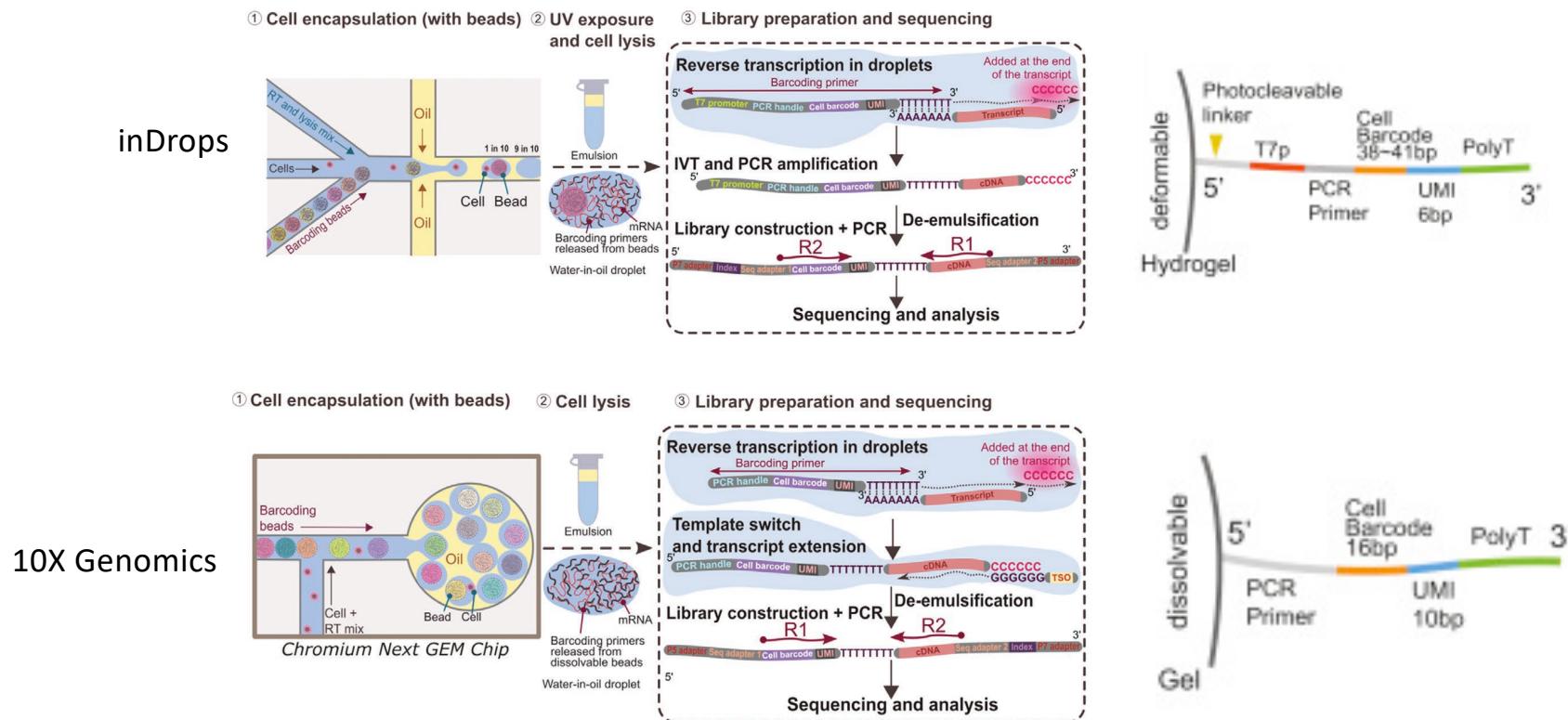
- Helps with the workflow
- Real-time cell count and viability
- True doublet rate
- Comprehensive statistics report

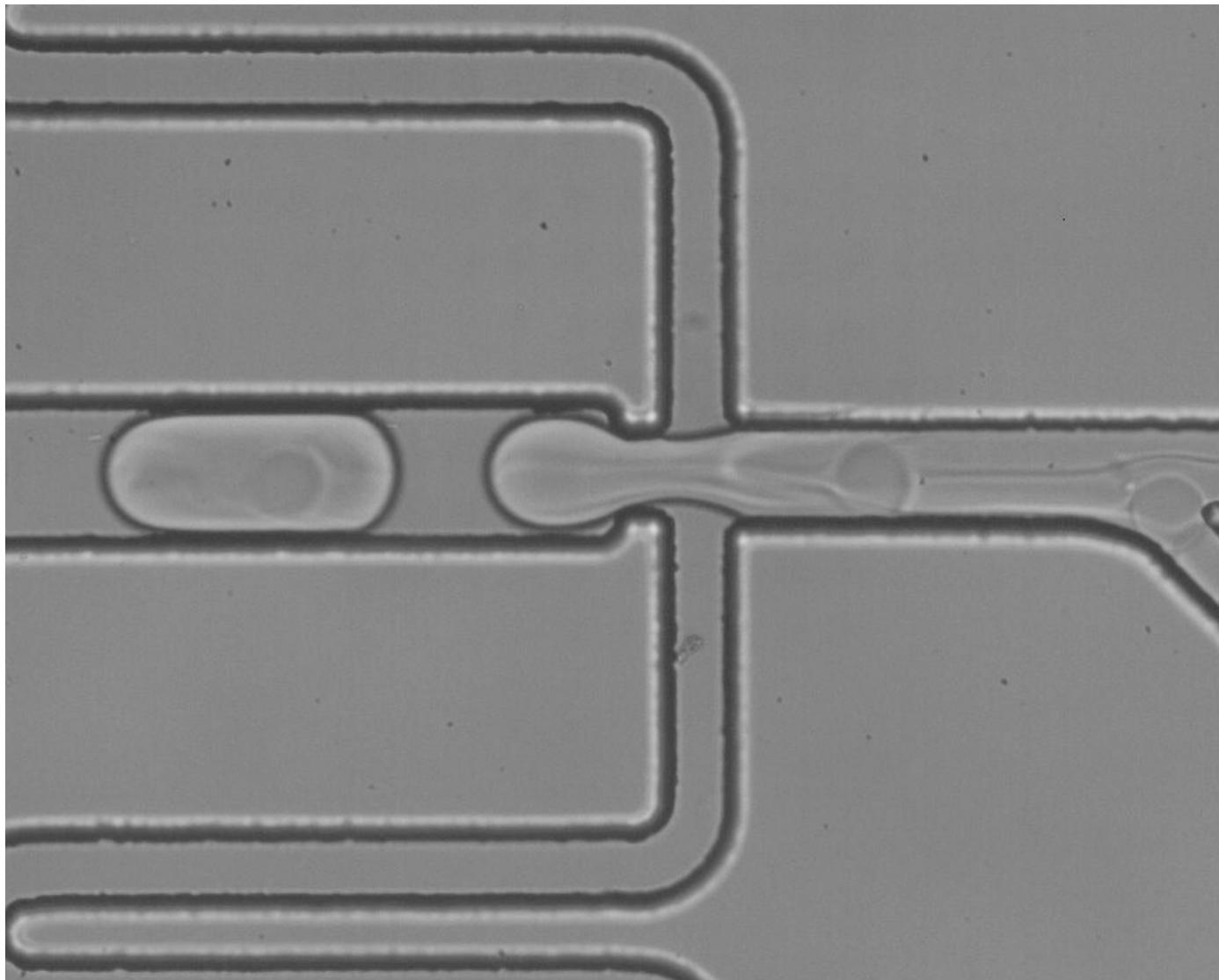
Rhapsody Express



- Manipulation of Microwell cartridge
- Portable/Benchtop
- Can be purchased individually

# Microfluidic-based droplet platforms



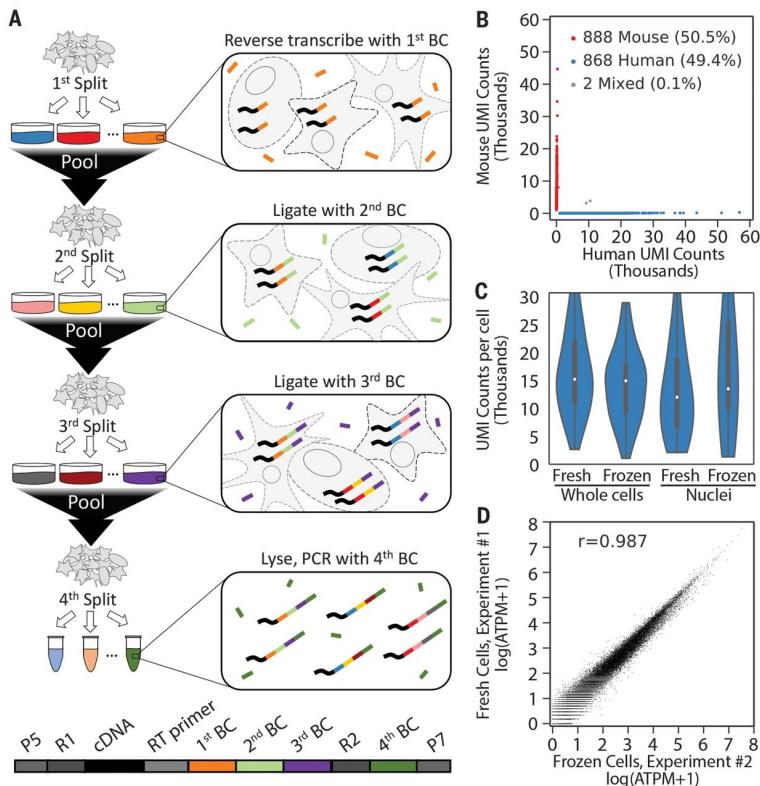


Video generated by  
Single Cell Core @ HMS

# scRNA-seq platforms that don't require specialized instrumentation

## Split-pool ligation-based transcriptome sequencing (SPLiT-seq)

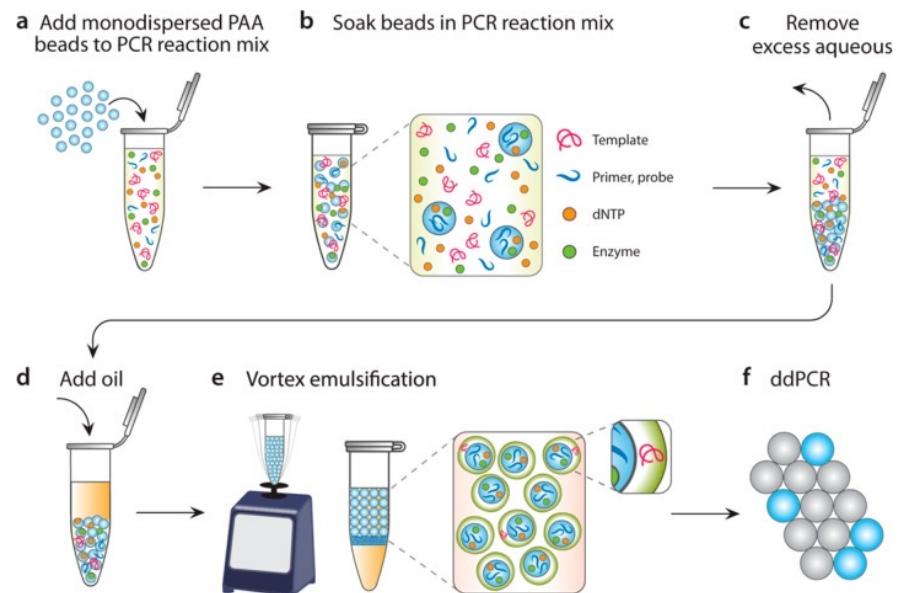
Commercialized by Parse Biosciences



Rosenberg et. al; SCIENCE, Vol. 360, Issue 6385, pp. 176-182

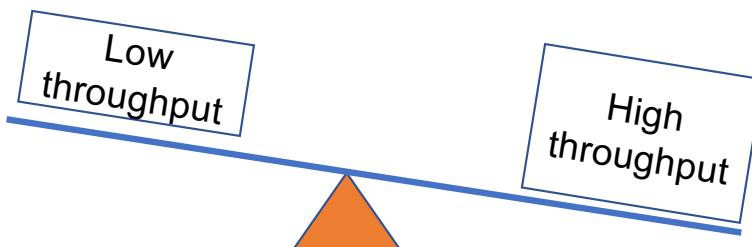
## Microfluidic-free droplet platform

Commercialized by Fluent Biosciences- Pre-templated Instant Partitions (PIP-seq)



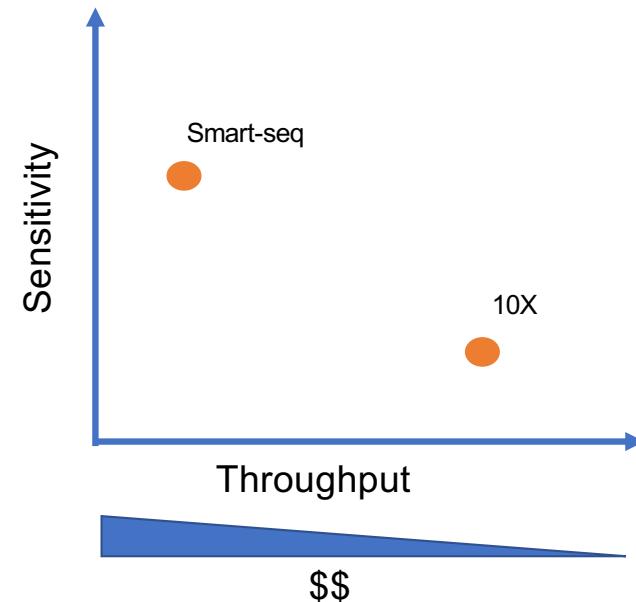
Hatori et. al; Anal Chem. 2018 Aug 21; 90(16): 9813–9820.

# Which platform should you choose?

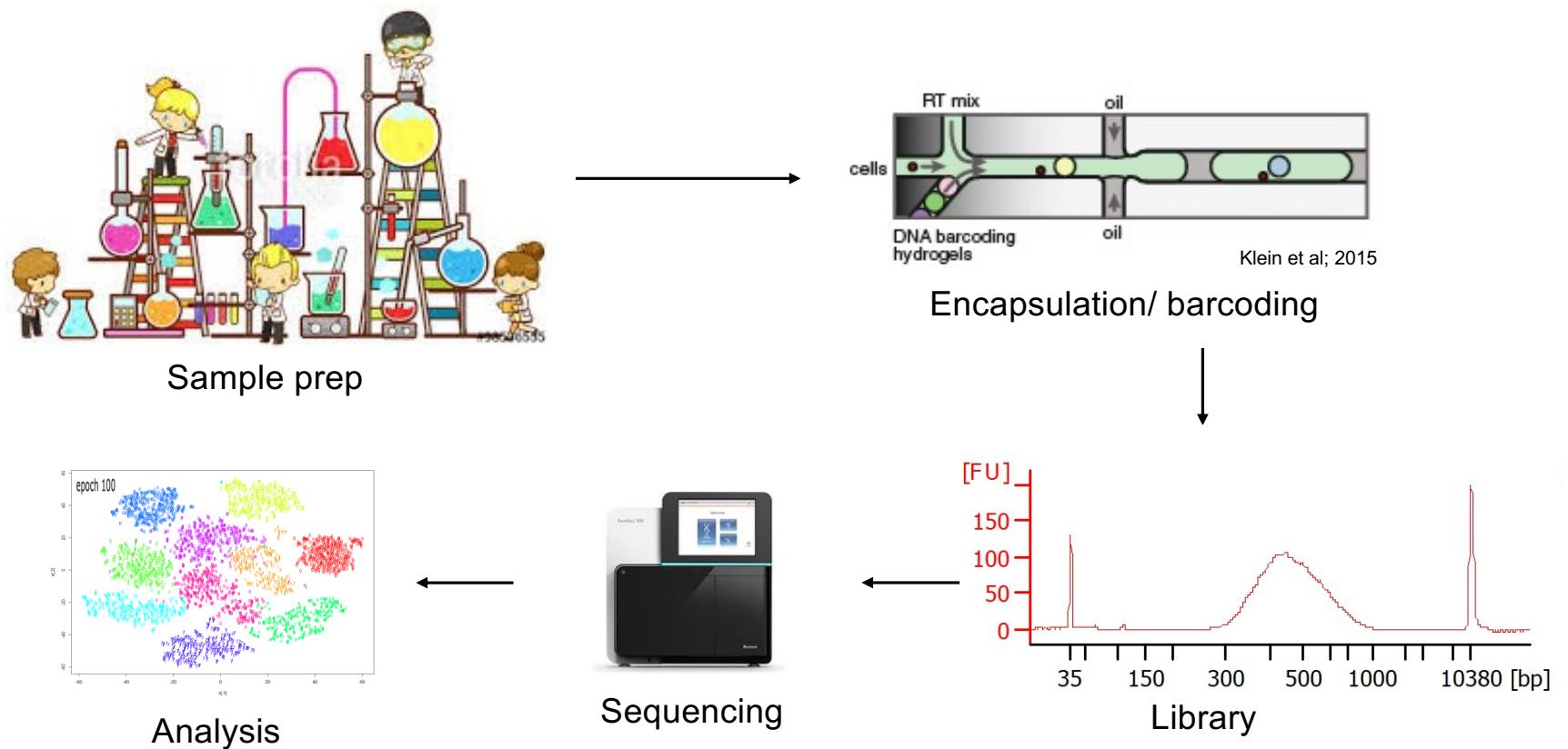


- Limited starting material, not enough for the high-throughput platforms.
- Interested in studying full length transcript, e.g. splice forms

- Need to profile a lot of cells- to detect a very rare sub-population in the tissue.
- Looking for overall heterogeneity in the tissue

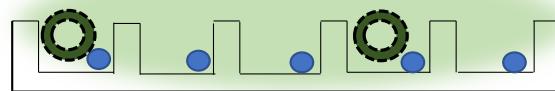


# Steps of barcoding mRNA in individual cell

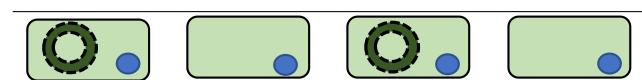


## High quality sample prep is a MUST regardless of the scRNA-seq platform

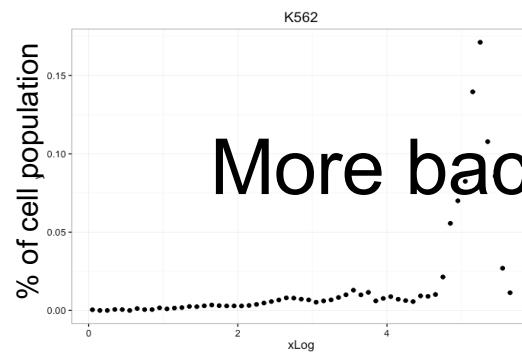
- High viability
  - More than 90% viability is ideal.
  - Cell membrane integrity is required until they are encapsulated.
- Good single cell suspension. No clumps.
- Clean prep with little or no debris.



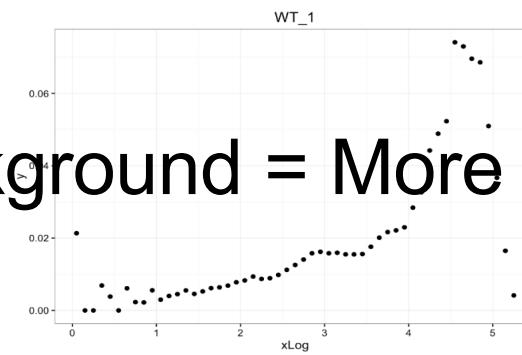
Nanowell-based barcoding platform



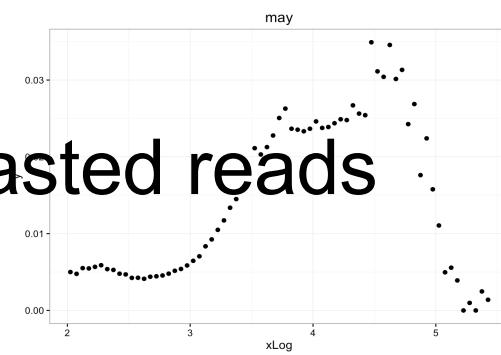
Droplet-based barcoding platform



Reads/ cell barcode  
Ideal data



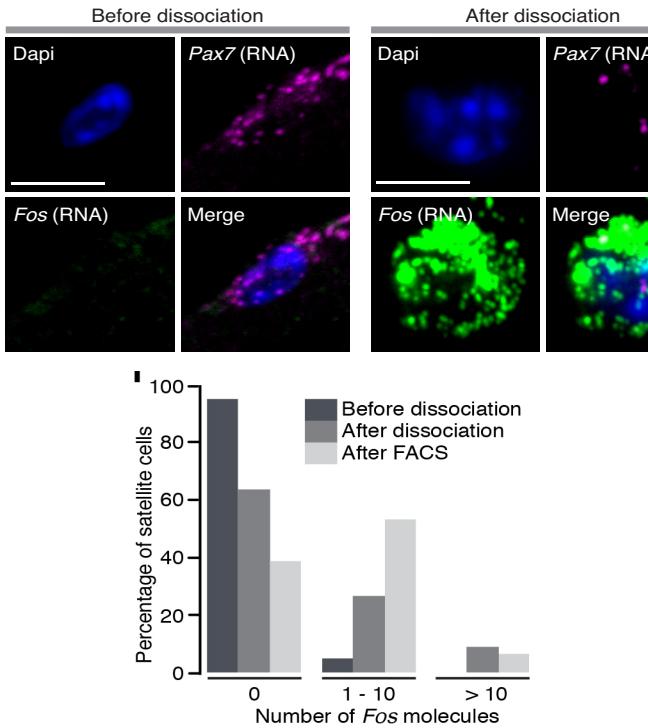
Reads/ cell barcode



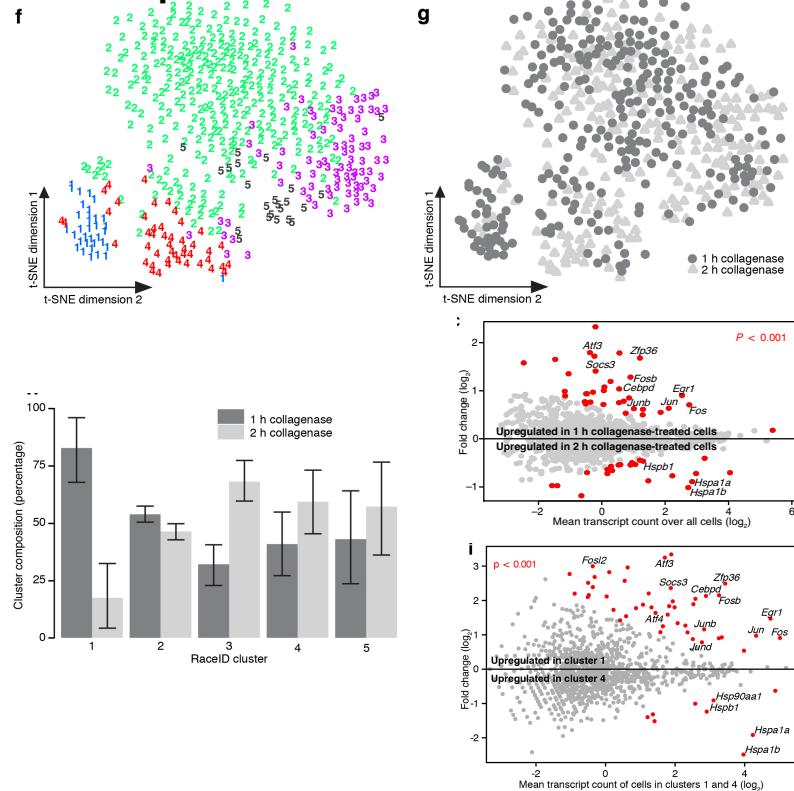
Reads/ cell barcode  
Free-floating RNA

# Transcriptome is stress-sensitive

**Single molecule FISH analysis shows that *Fos* expression is induced during the SC isolation procedure.**

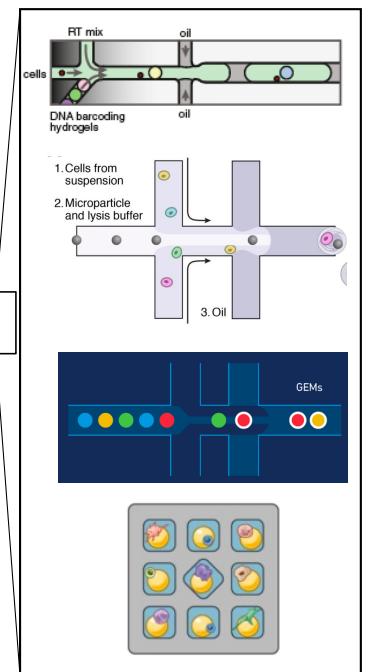
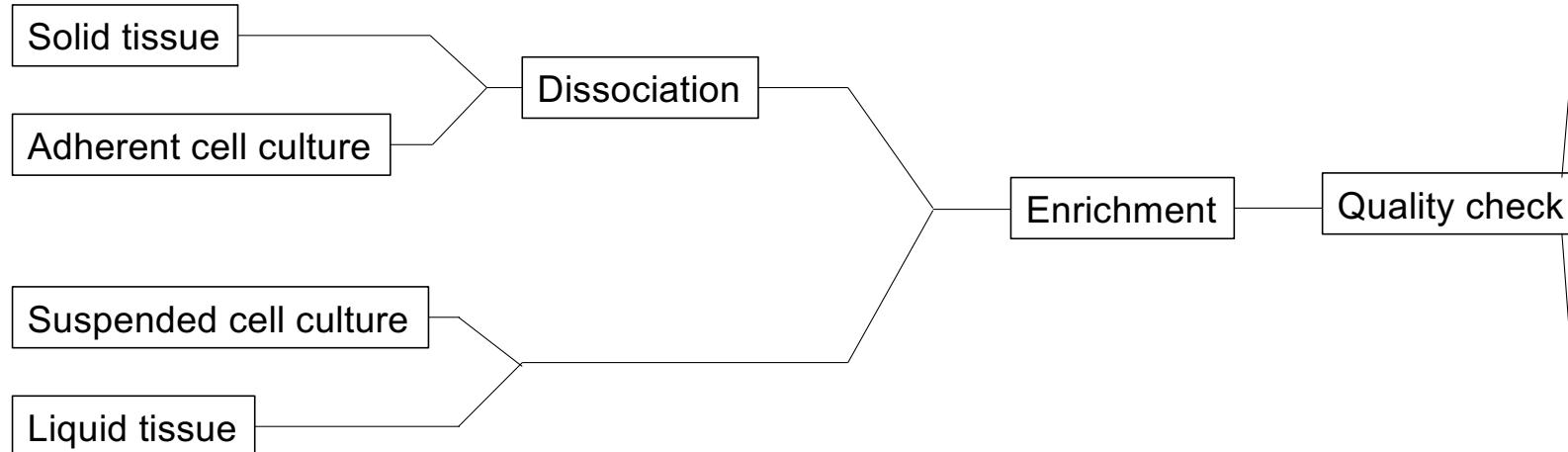


**Dissociation time experiment confirms that the dissociation procedure influences the transcriptome of SCs.**

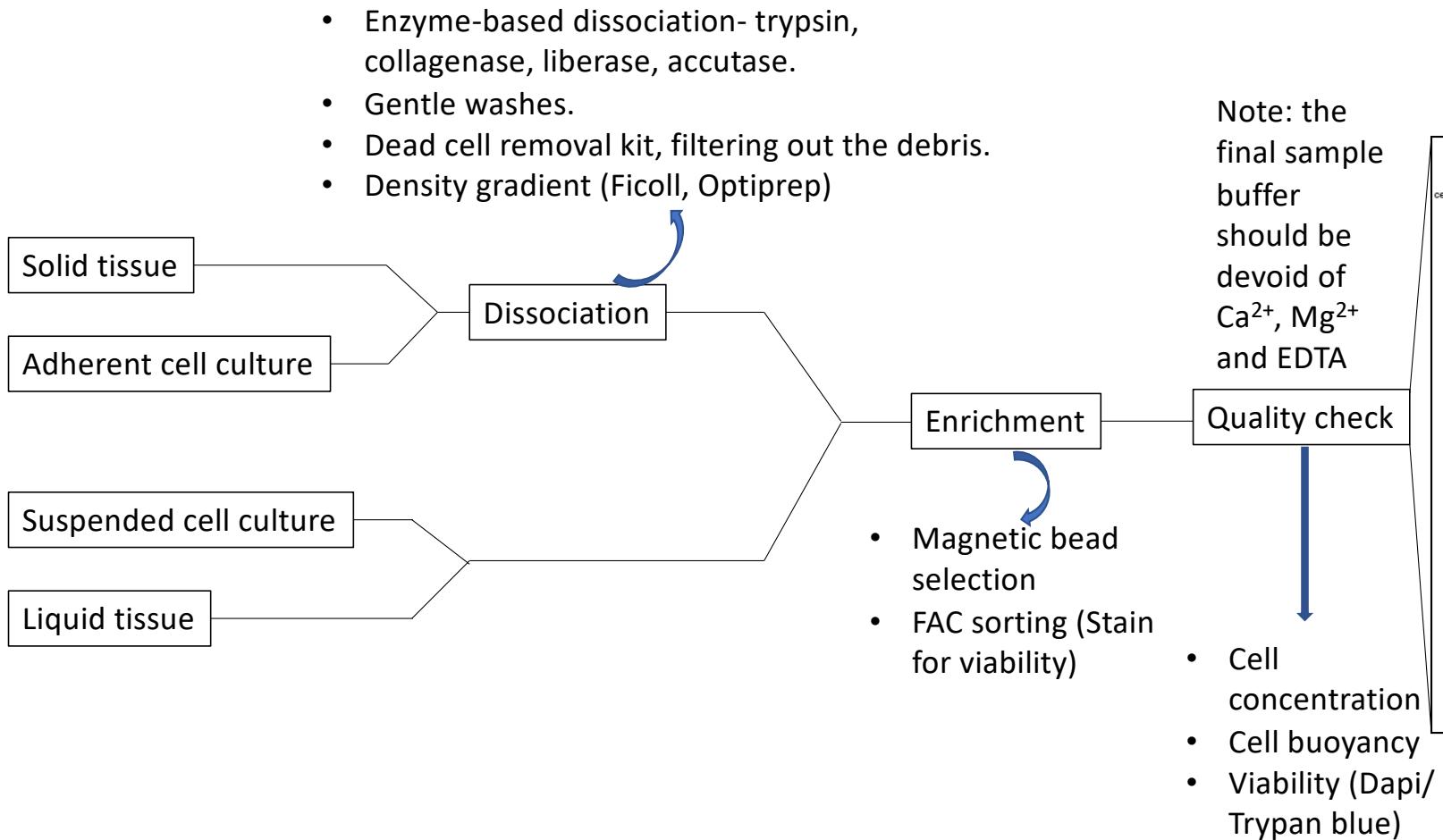


# Sample prep

- ❖ What is your sample of interest? And how would you obtain that?
  - Which population in a tissue should be examined?
  - Does it require some sort of enrichment?



# Sample prep protocol varies from cell-type to cell-type



# Enrichment methods: Pros and cons

## FACS (Fluorescence activated cell sorting)

### Pros:

- Enrichment is robust. Can be really useful for rare population of cells.
- Yields good single cell suspension.
- Live/dead sorting by DNA stains, eg. DAPI.

### Cons:

- Uses high pressure to sort the cells, therefore can be pretty harsh.
- Can introduce bias in the experiment. Using a broad marker is recommended.
- Long sample prep protocol.

## MACS (Magnetic-activated cell sorting)

### Pros:

- Gentle on cells.
- Fast protocol.
- Greater number of cells can be processed at a time.
- Not limited by FAC sorter availability.

### Cons:

- Number of available surface marker-conjugated to magnet is limited.
- Enrichment is not precise. Not applicable for rare population.

# Best practices to obtain high quality sample

- Optimize a dissociation protocol that is best-suited for your cell type of interest.
- Run a pilot experiment. Do not rush to the final experiment.
- Short sample prep time.
- Maintaining low temperature.
- Gentle treatment-
  - gentle lysis condition (low temp, short time)
  - short FACS time, slow sorting, bigger nozzle (in certain cases)
  - Gentle centrifugation (300-500xg) and resuspension
  - Removing debris by filter or density medium
- Include BSA (upto 1%) or FBS (upto 2%).

# Alternatives to freshly harvested samples

- Challenges with freshly harvested samples
  - Sample prep is time-consuming and time-sensitive.
  - Limits the number of samples that can be run a day.
- Fixed and cryopreserved samples
  - Methanol fixation -> rehydration with SSC or PBS(RNase inhibitor, maybe)
  - 5% DMSO as a cryoprotectant is compatible with scRNA-seq
- Single nuclear samples
  - Entire tissue can be frozen and stored.

# scRNA-seq vs snRNA-seq

## snRNA-seq

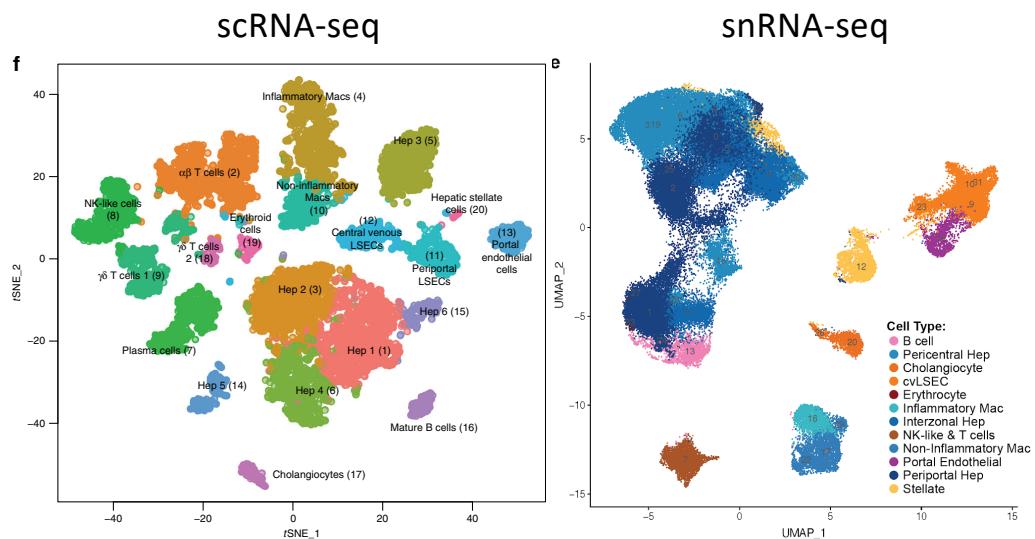
### Advantages:

- Helps overcome the limitation associated with tissue dissociation
- Enables the analysis of biobanked samples at a single cell resolution
- Helps simplify the experimental design involving many samples or helps mitigate the challenges with time-course experiment
- The preparation protocol is often faster than single cell prep

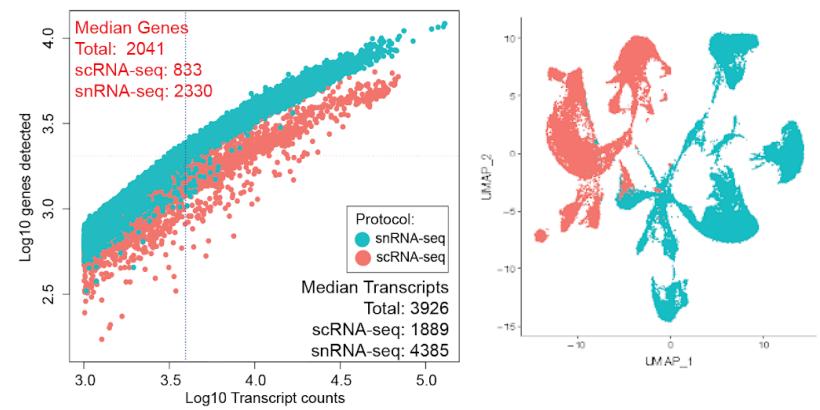
### Disadvantages:

- The material per single nuclei is much less compared to that from single cell.
- The nuclei can be leaky, and the mRNA might escape, which in turn can increase the background.
- Single nuclei prep has a lot of debris.
- Counting nuclei is not trivial.

# scRNA-seq vs snRNA-seq: Liver tissue



- All major hepatic cell types were represented in both scRNA-seq and snRNA-seq, but were captured at different frequencies.
- Cholangiocytes and parenchymal cells were underrepresented.
- Immune cells were more easily detected in the single cell prep.



- SnRNA-seq captured a greater diversity of genes than scRNA-seq- high proportion of UMIs in scRNA-seq data are derived from transcripts encoding ribosomal proteins and genes encoded in the mitochondrial genome, which are not present in snRNA-seq data
- scRNA and snRNA don't cluster together due to the systemic difference between the RNA found in the nucleus vs in the cytoplasm. Also the preparation methods for these two are different.

# Cryopreservation vs Fixation

- Freshly prepared samples are the best.
- Cryopreservation
  - 5-10% DMSO has been shown to be the most effective cryoprotectant and compatible with scRNA-seq
  - During the thawing process some cell loss is inevitable. Transcriptomes are mostly comparable (sample-specific)
  - One sub-population of cells might be more susceptible to freeze-thaw than the others
- Fixation
  - 100% Methanol (10X supported) or dithio-bis-succinimidyl propionate (DSP), also called Lomant's Reagent
  - Fixation is harsh on the cell membrane that might result into high cell lysis and cell loss.
  - 100% methanol-fixed samples are rehydrated using SSC

## How many cells do I need to analyze? How deep do I need to sequence?

The answer is not simple!

- Depends on the question.
- Rule of thumb: 50-100 cells with unique transcriptome signature is necessary for forming a distinct cluster in a t-SNE plot.
- Rare cell type of your interest- need to analyze many cells.
- For overall heterogeneity- fewer cells might be enough.
- Sequencing depth- rarer the transcript, higher should be the sequencing depth.

# Challenge: Batch effect

- Many sources
  - Sample prep
  - Animals from different clutches/ litters
  - Different batches of reagents
  - Operator
  - Library prep performed at different time
  - Different sequencing run
- Solution
  - Wet-lab solution
    - Use biological replicates. Standard statistical considerations apply. This can become expensive.
    - Same reagent batch, if possible
    - Keeping the operator same
    - Prepping libraries and sequencing together
    - Pooling of samples by sample barcoding (hashtagging/ MULTI-seq)
  - Dry-lab solution
- Sometimes batch effects are unavoidable- patient samples, time-course experiments, perturbation experiments
- Batch effect correction works best when individual samples contain sufficient internal complexity to identify shared sources of transcriptional variation

# Thank you!

