Introduction to Single Cell RNA Sequencing

Sarah Boswell

Director of the Single Cell Core, Harvard Medical School

Director of Sequencing Technologies, Laboratory of Systems Pharmacology

Staff Scientist/Sequencing Specialist, Systems Biology, Springer Lab

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

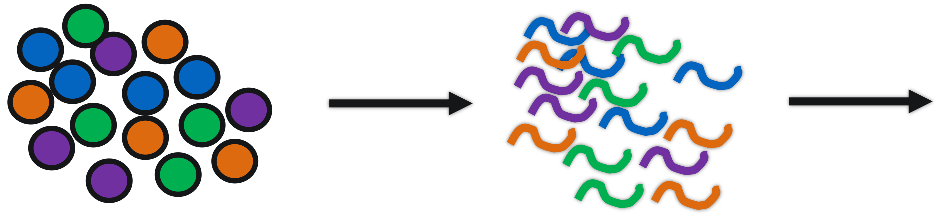
sarah_boswell@hms.harvard.edu

Introduction to Single Cell RNA Sequencing

- Common applications of single cell RNA sequencing.
- Overview of single cell RNA sequencing platforms.
- Modified scRNA-seq workflows
- Sample preparation and experimental design.
- Effects of sample prep and sample type on analysis.

Bulk vs Single Cell RNA Sequencing (scRNA-seq)

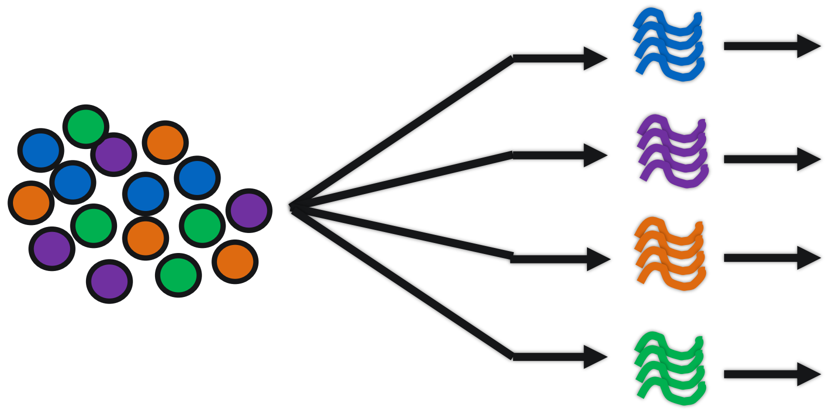
Bulk RNA-seq



average expression level

- comparative transcriptomics
- disease biomarker
- homogenous systems

scRNA-seq



Population 1

Population 2

Population 3

Population 4

- define heterogeneity
- identify rare cell population
- cell population dynamics

Transcriptome Coverage (mRNA)

1. mRNA: TruSeq RNA-Seq (gold standard)
    - ~20,000 transcripts
        - More when consider splice variants / isoforms
    - Observe 80-95% of transcripts depending on sequencing depth

2. Low Input Methods
    - 4000-6000 transcripts per sample
        - Limiting to transcripts observed across all samples
    - Observe 20-60% of the transcriptome

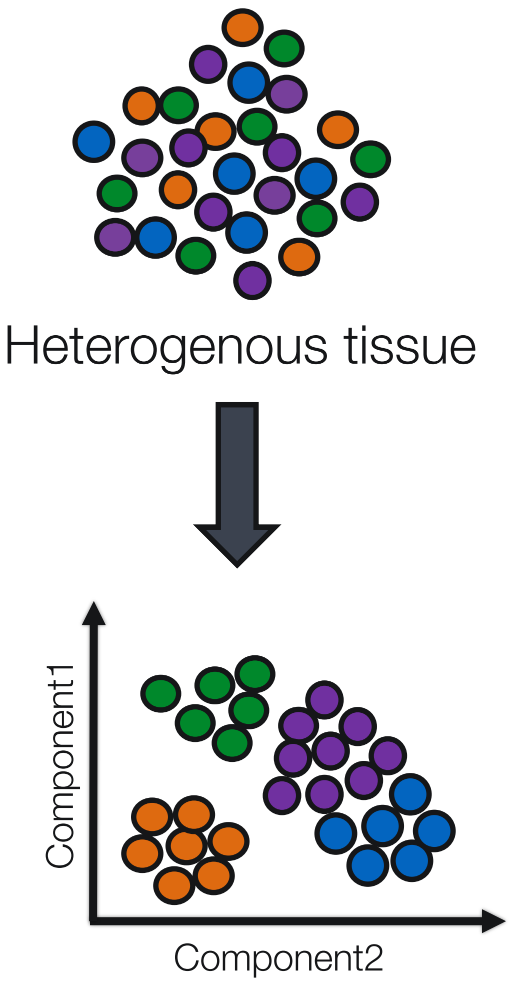
3. Single Cell Methods
    - 200 -10,000 transcripts per cell
    - Observe 10-50% of the transcriptome
    - Many transcripts will show up with zero counts in every cell (eg. GAPDH, ACTB).
    - If you only looked at transcripts observed in all cells numbers drop dramatically.

The World Between Bulk & scRNA-seq

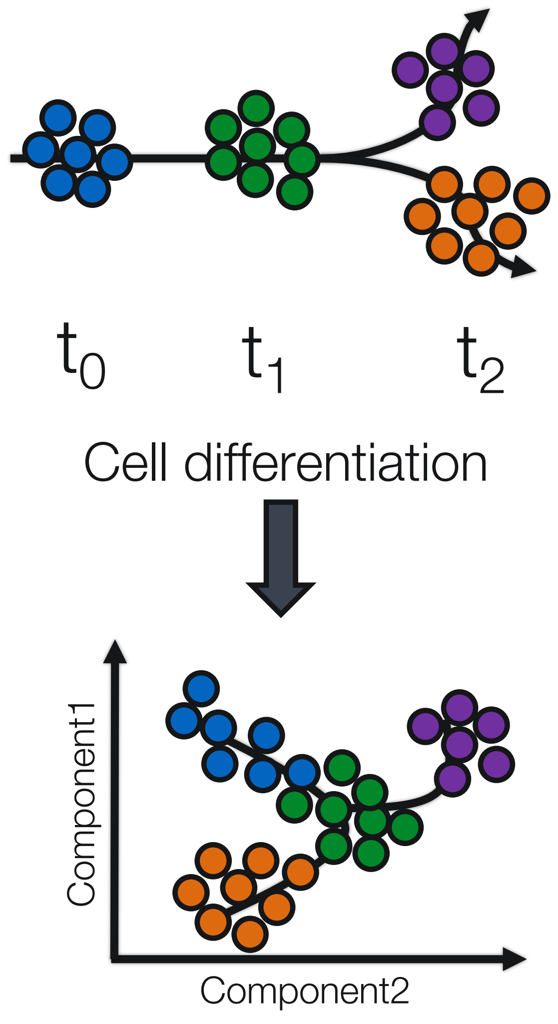
	Deep RNA-seq	Sort-seq	Low input	scRNA-seq
Transcriptome Coverage	High	High	Moderate	Low
Throughput	Moderate	Low	High	Low
Cell Subtype Information	None	Moderate	None	High
Sequencing Depth	Moderate	Moderate	Low	High
Cost per Sample	Moderate	Moderate	Low	High

Common Applications of scRNA-seq

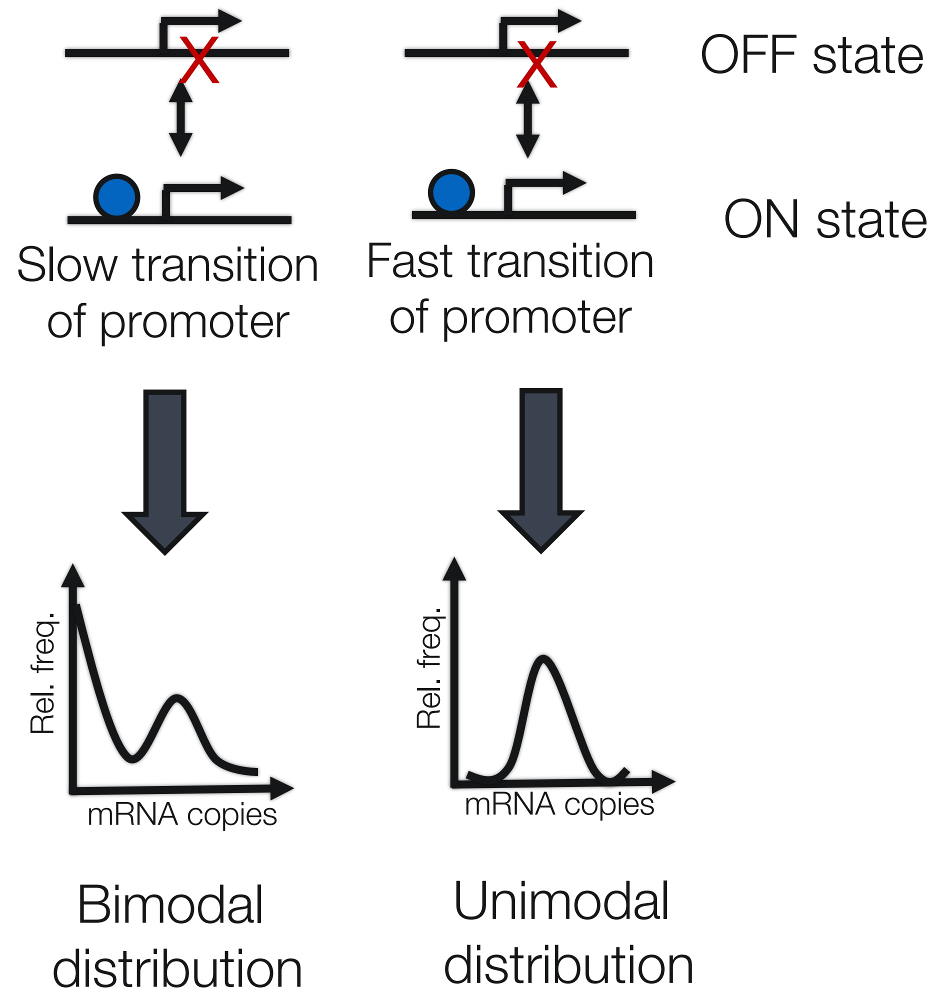
Studying heterogeneity



Lineage tracing study

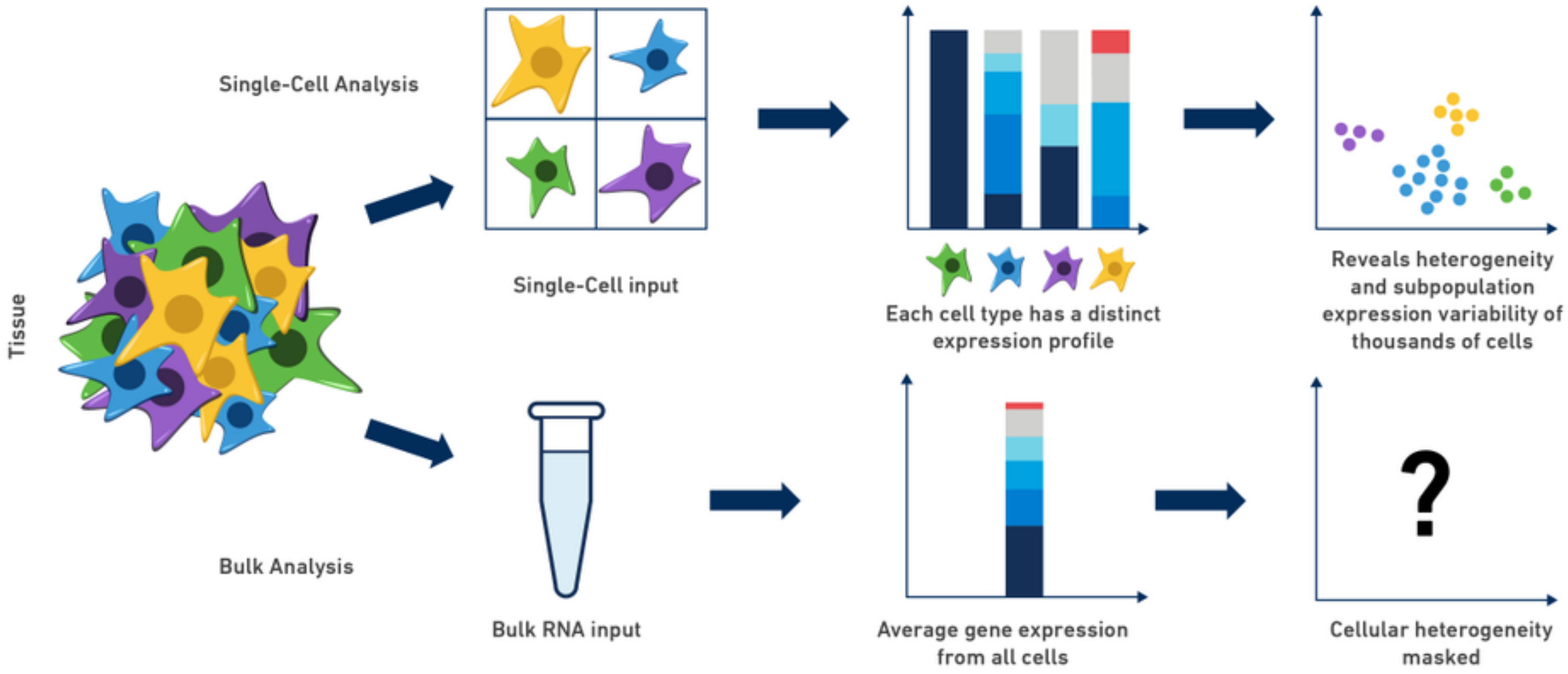


Stochastic gene expression



Liu S and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges, F1000 Research 2016 (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 2014 (doi: 10.1016/j.cell.2014.02.010)

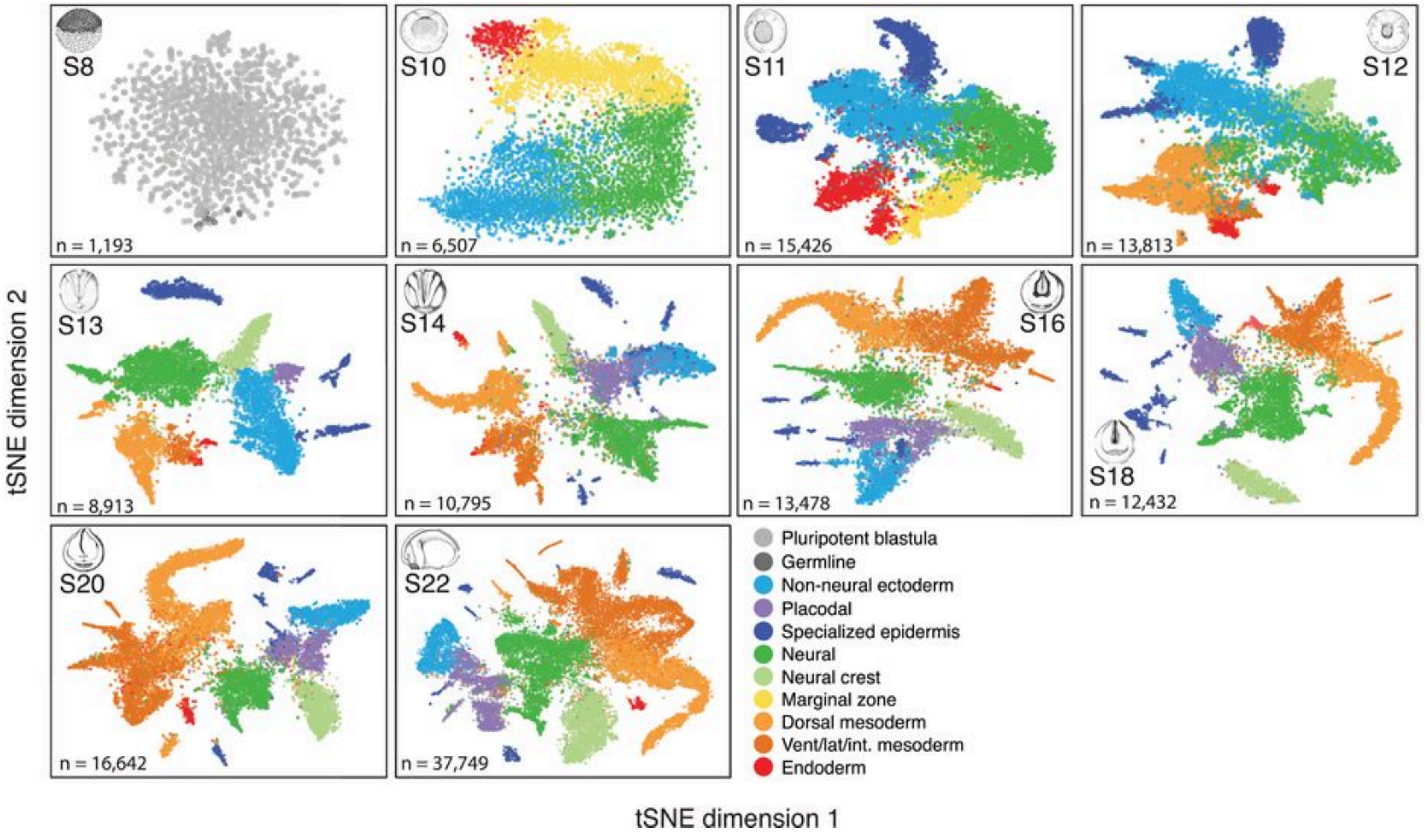
Tumor, Tissue, Organoid Heterogeneity



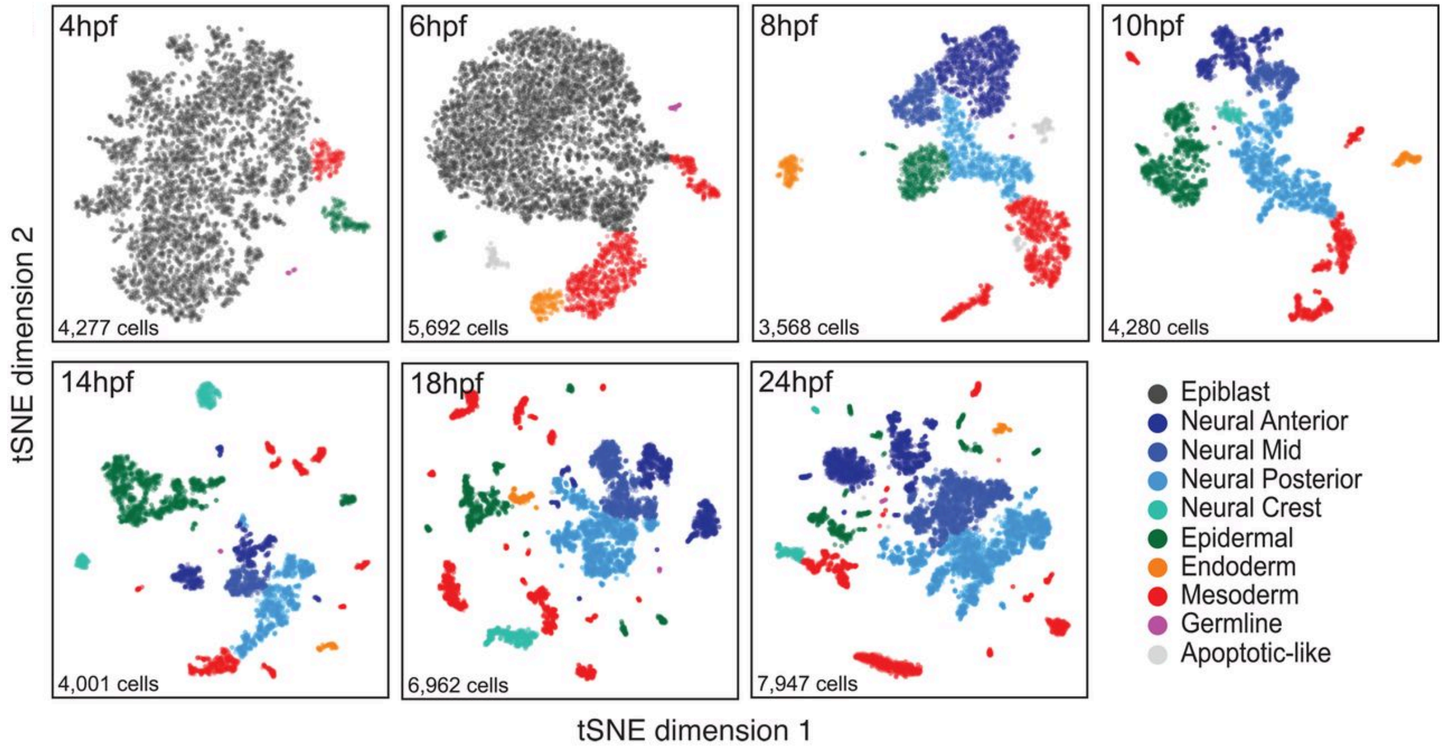
https://community.10xgenomics.com/t5/10x-Blog/Single-Cell-RNA-Seq-An-Introductory-Overview-and-Tools-for/ba-p/547

Development Lineage Tracing

Frog

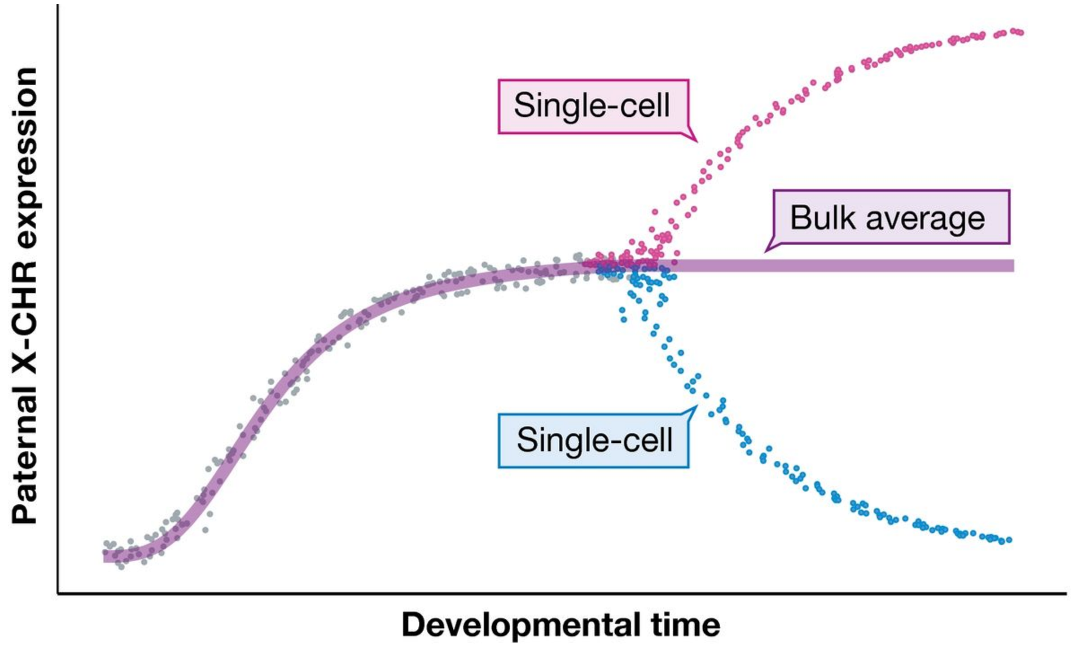


Zebrafish

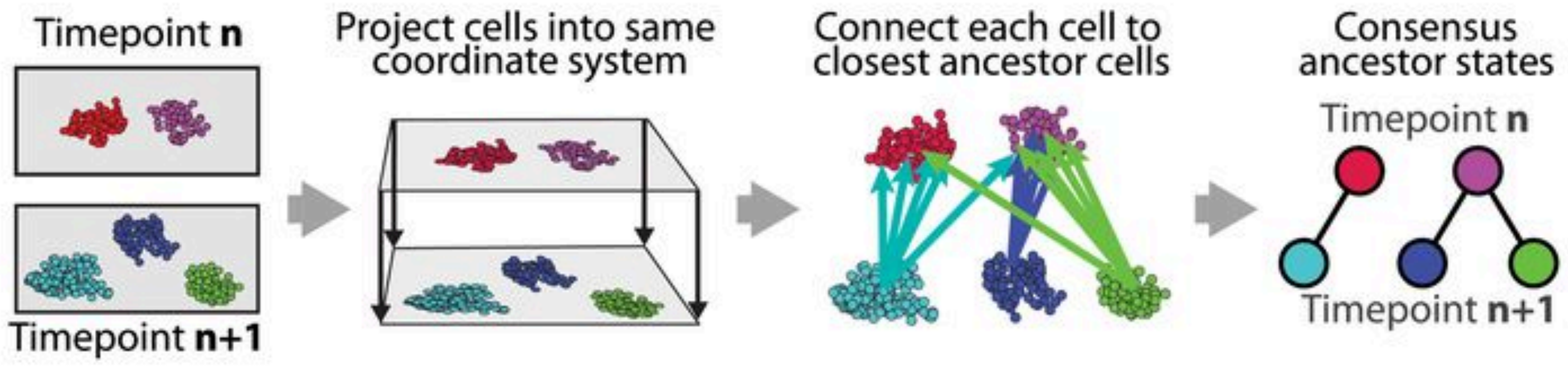


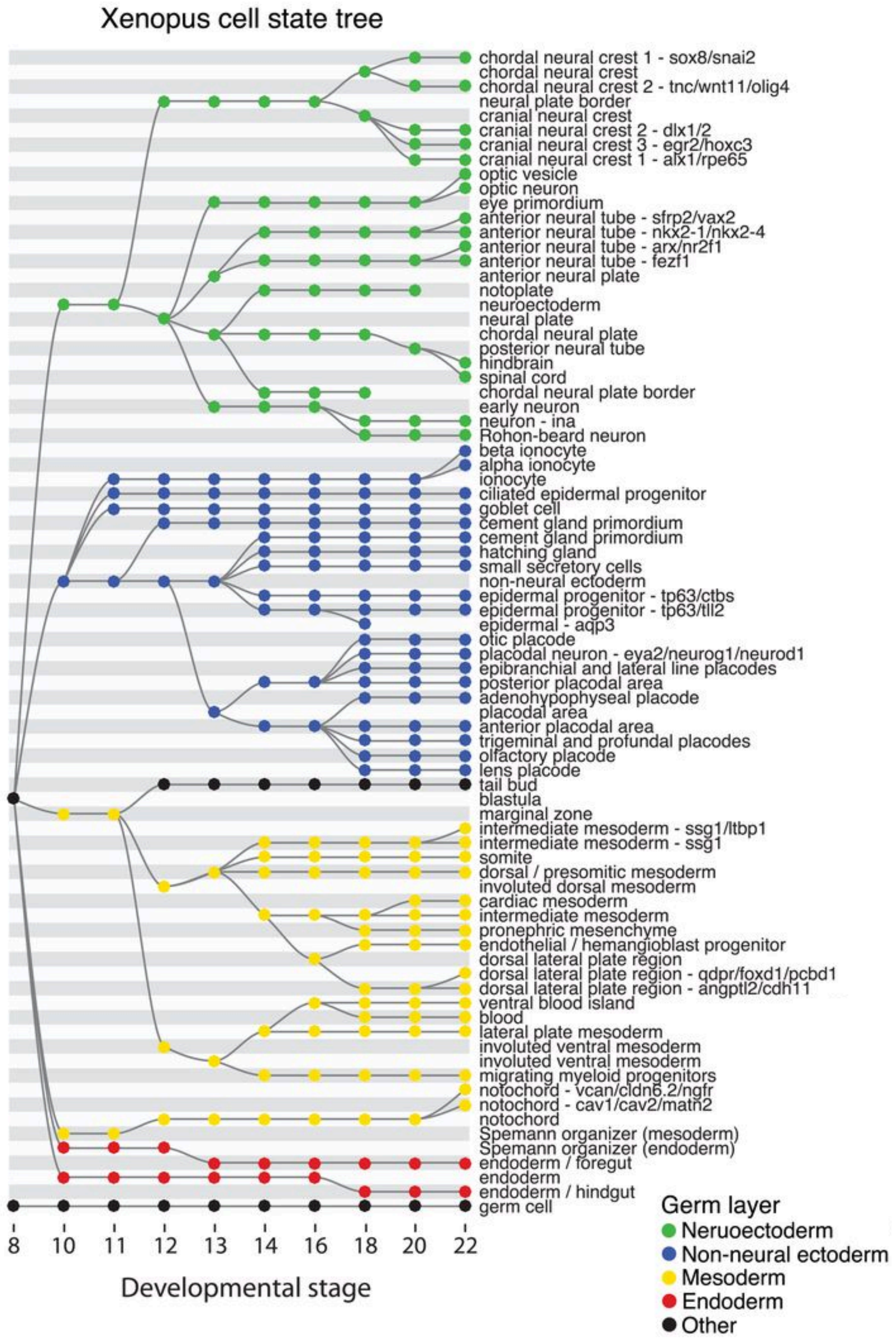
JA. Briggs et al. The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution, Science 01 Jun 2018 (DOI: 10.1126/science.aar5780)
DE Wagner et al. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo, Science 01 Jun 2018 (DOI: 10.1126/science.aar4362)

Development Lineage Tracing



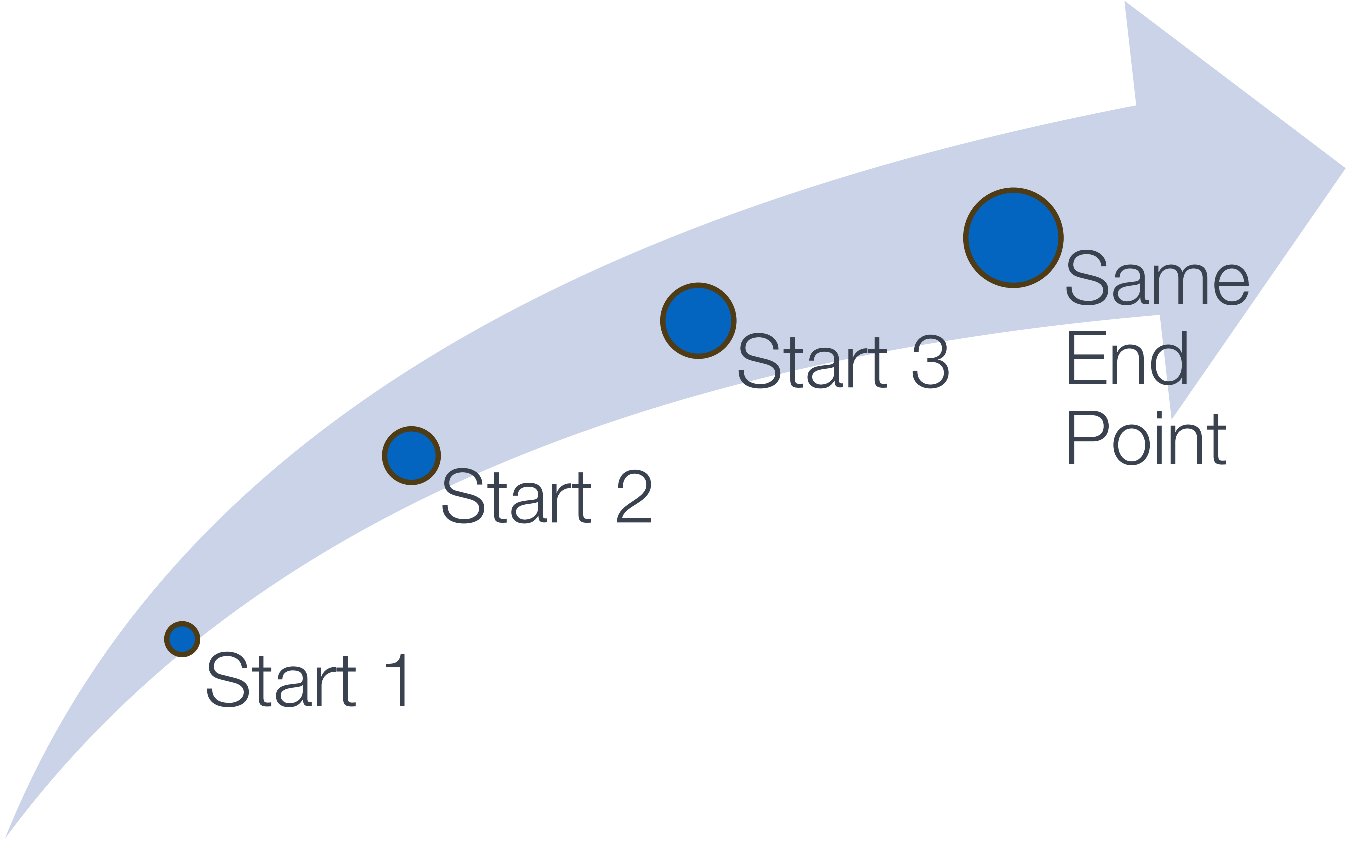
© EMBO

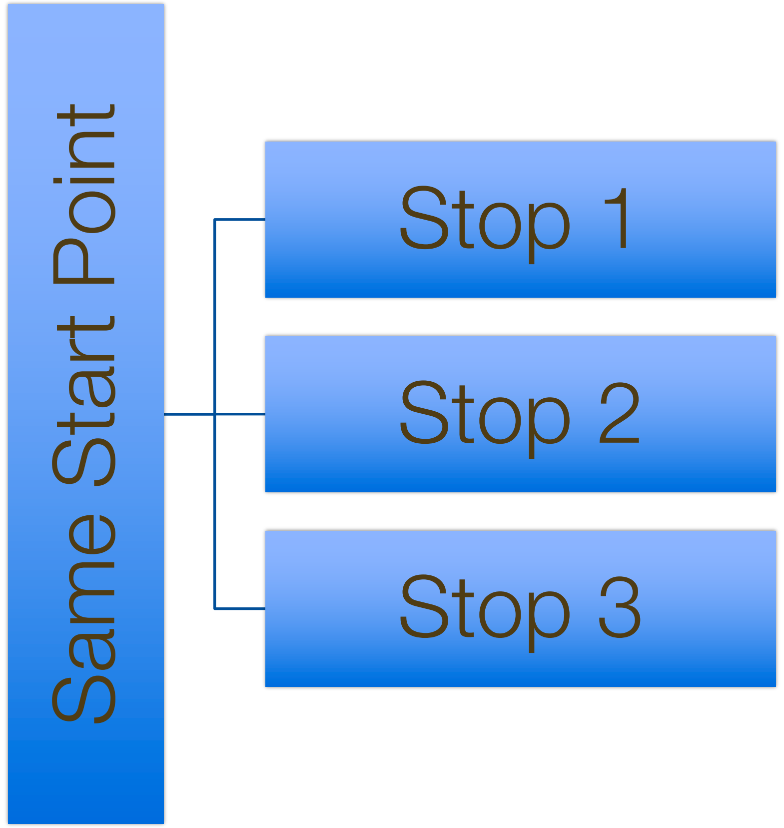




JA. Briggs et al. The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution, Science 01 Jun 2018 (DOI: 10.1126/science.aar5780)
JA. Griffiths et al. Using single-cell genomics to understand developmental processes and cell fate decisions, MSB (2018) (DOI 10.15252/msb.20178046)

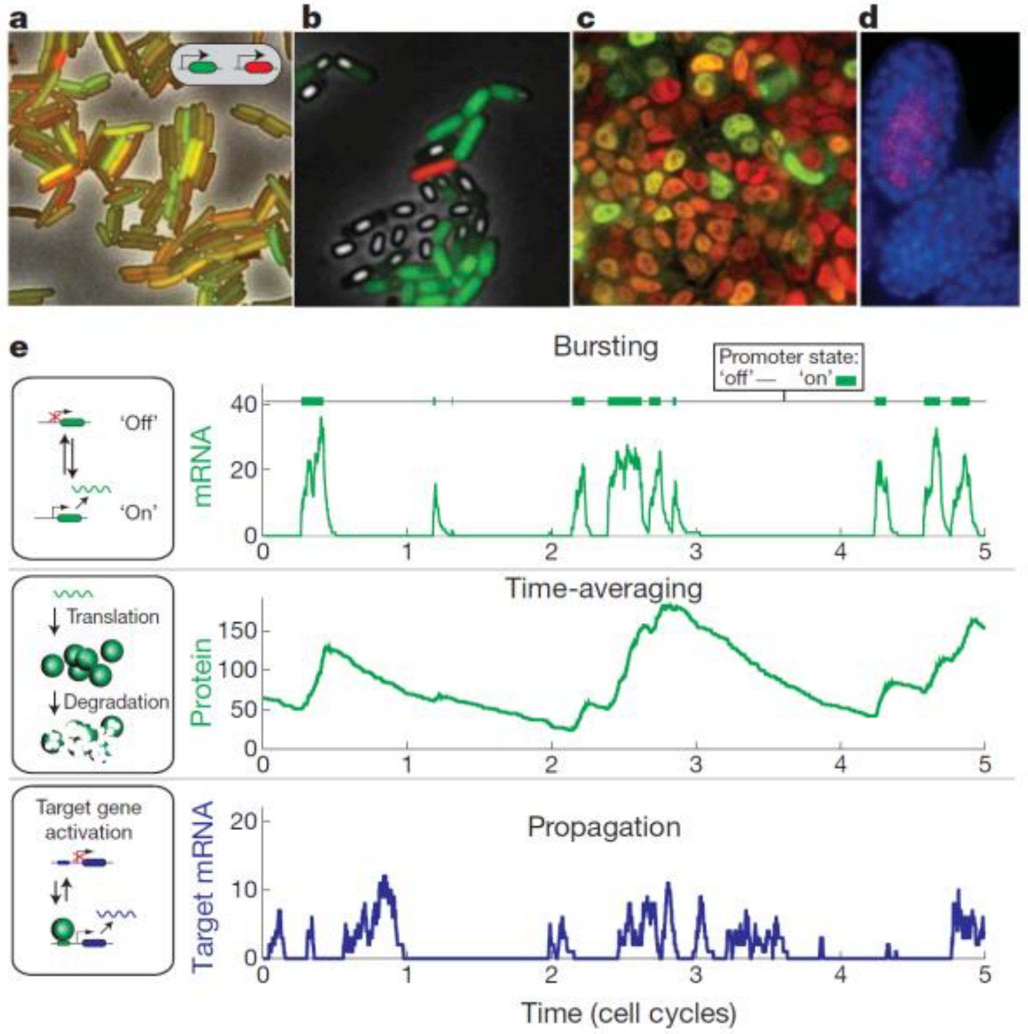
Time Course or Development Experiment





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

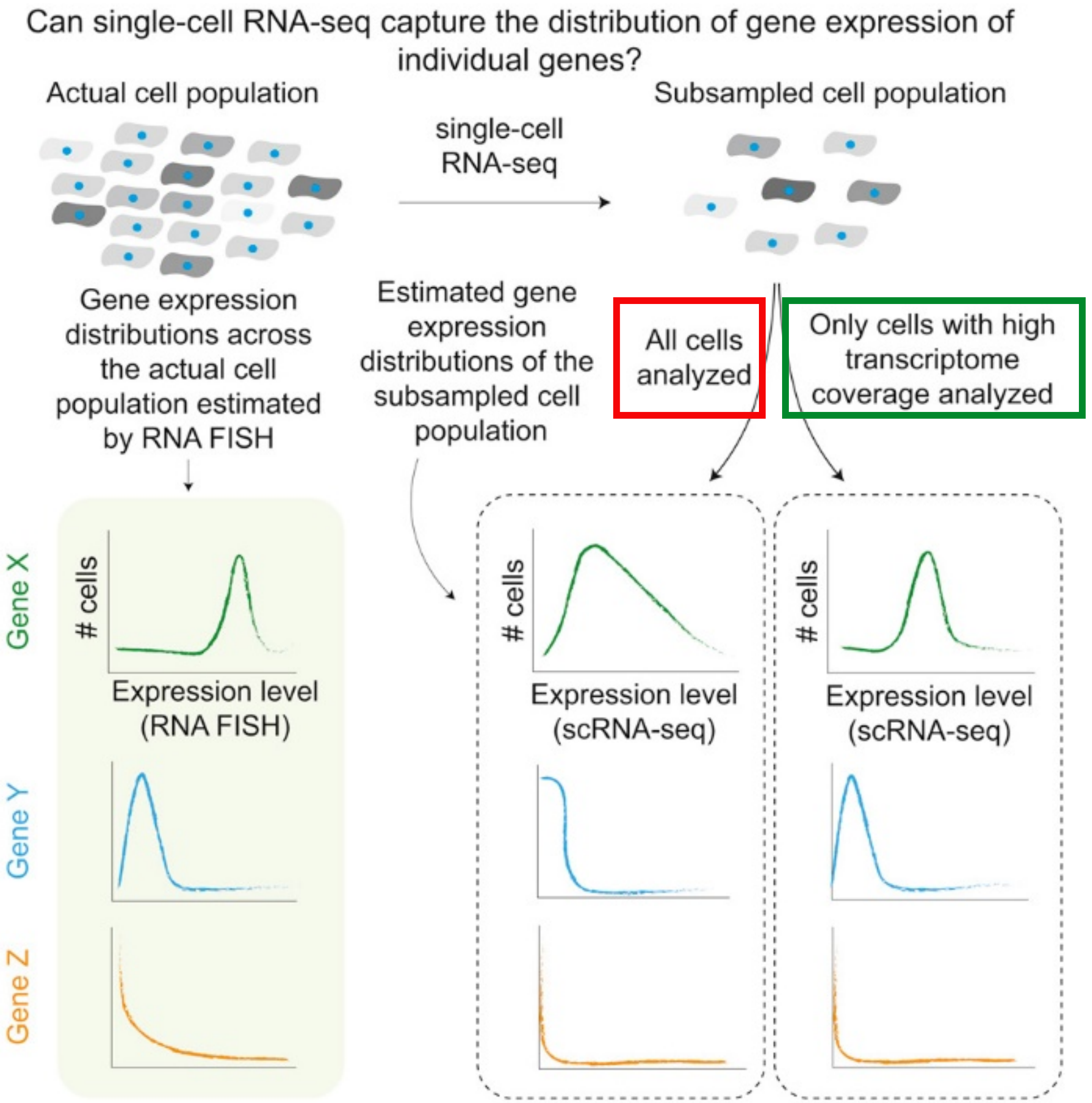
Stochastic Gene Expression



- Gene expression is heterogeneous and "bursty".
- Genes fluctuate between "On" and "Off" promoter states.
- Stochastic expression of one gene can propagate to generate more stochasticity in downstream genes.

Eldar & Elowitz; Functional roles for noise in genetic circuits, Nature 2010 (doi: 10.1038/nature09326)
B Hwang et al. Single-cell RNA sequencing technologies and bioinformatics pipelines, EMM 07 Aug 2018 (doi: 10.1038/s12276-018-0071-8)

Stochastic Gene Expression



- Low mRNA capture efficiency of scRNA-seq makes it difficult to draw definitive conclusions about expression at the single-cell level.

- Number of cells and depth of sequencing critical for understanding rare gene expression phenotypes.

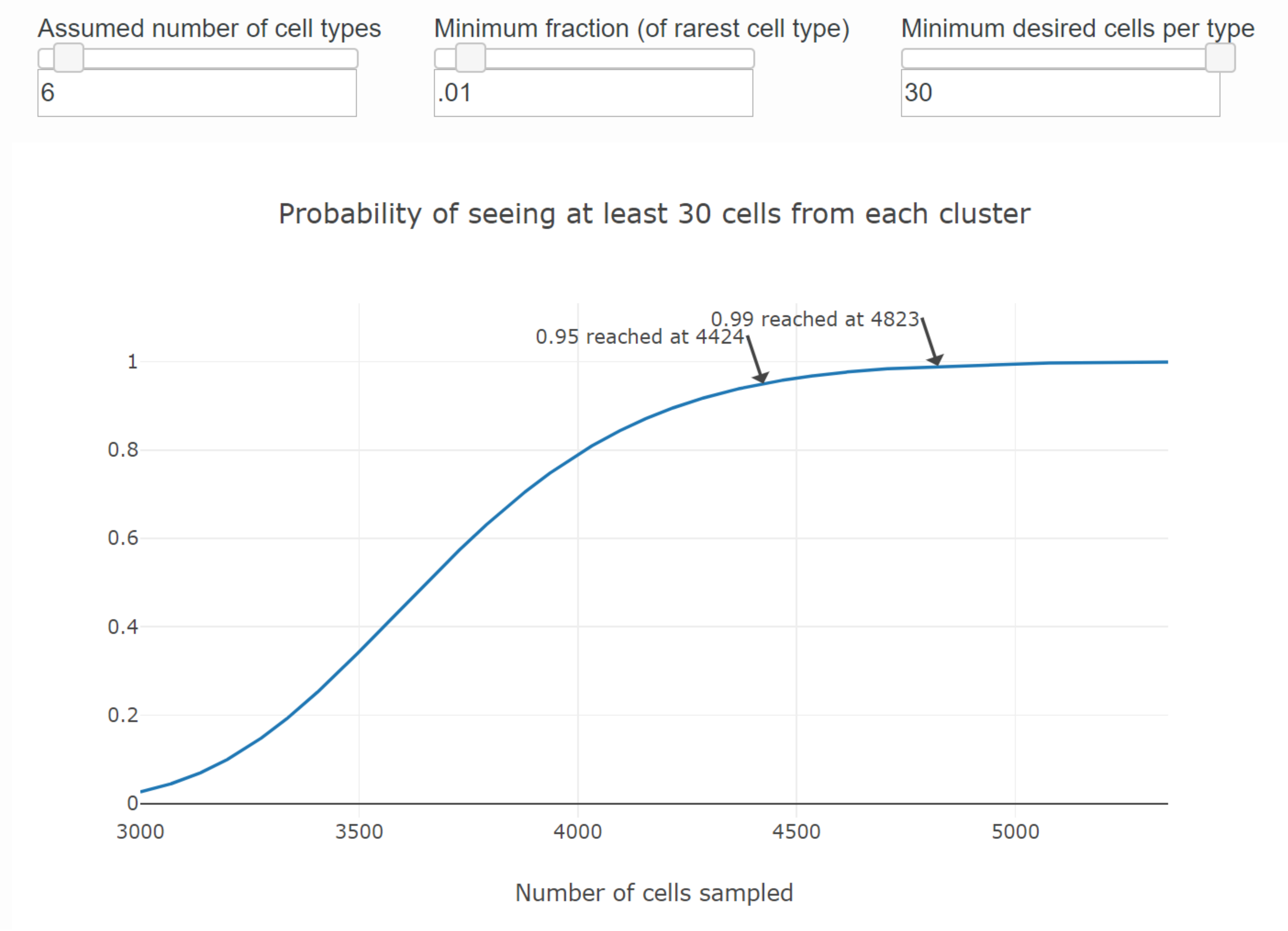
E Torre et al. Rare Cell Detection by Single-Cell RNA Sequencing as Guided by Single-Molecule RNA FISH, Cell Systems 28 Feb 2018 (DOI: 10.1016/j.cels.2018.01.014)

More Cells or More Sequencing Reads?

- Required number of cells increases with complexity of the sample.
- As the number of genes involved in the biology decrease then the coverage requirements increase (more reads).
- Cell-type classification of a mixed population usually requires lower read depth and can be sequenced at 10,000-50,000 reads per cell.
- Suggest starting with 25,000-55,000 reads per cell. Can always re-sequence your samples.

A. Hague et al. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications, Genome Med 2017 (DOI: 10.1186/s13073-017-0467-4)
JA. Griffiths et al. Using single-cell genomics to understand developmental processes and cell fate decisions, MSB (2018) (DOI 10.15252/msb.20178046)

https://satijalab.org/howmanycells



Assumed number of cell types
6

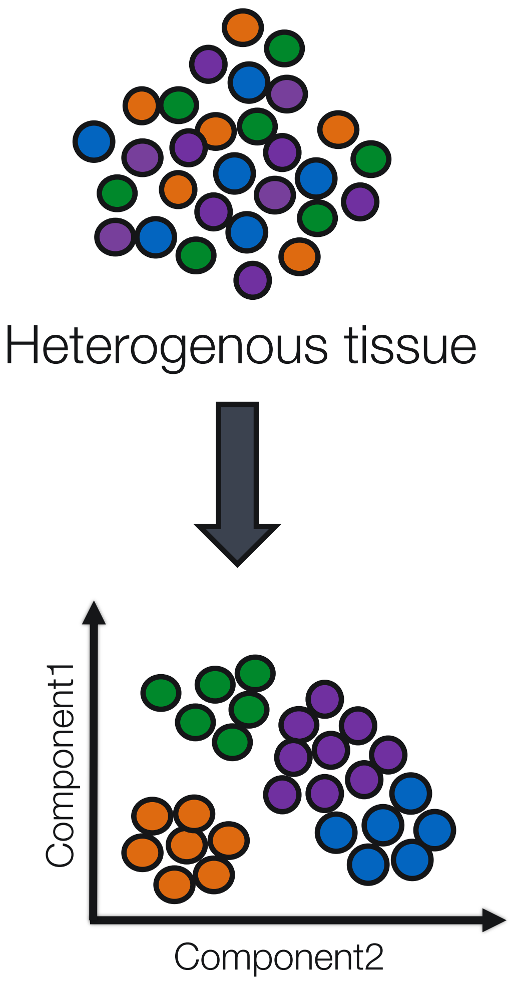
Minimum fraction (of rarest cell type)
.01

Minimum desired cells per type
30

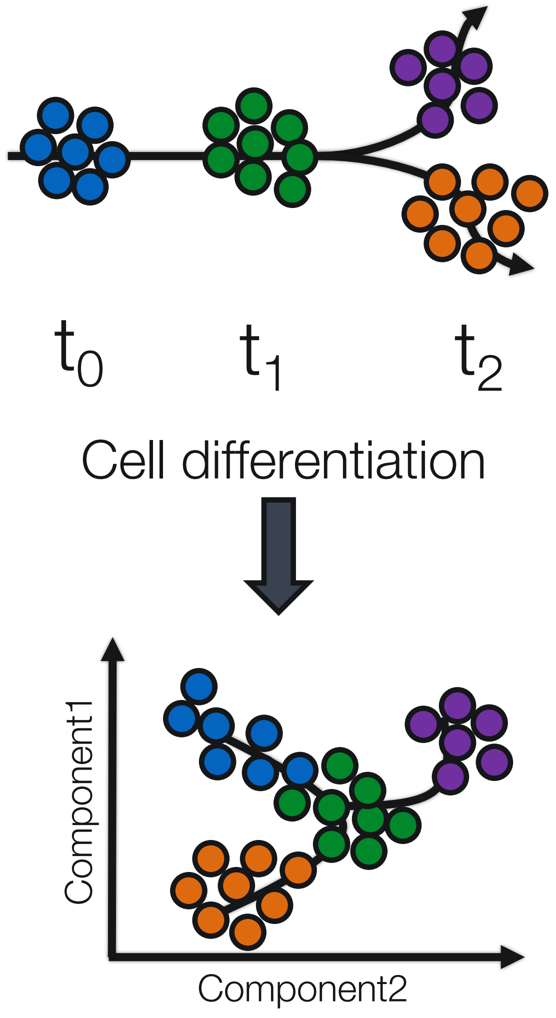
Probability of seeing at least 30 cells from each cluster

Common Applications of scRNA-seq

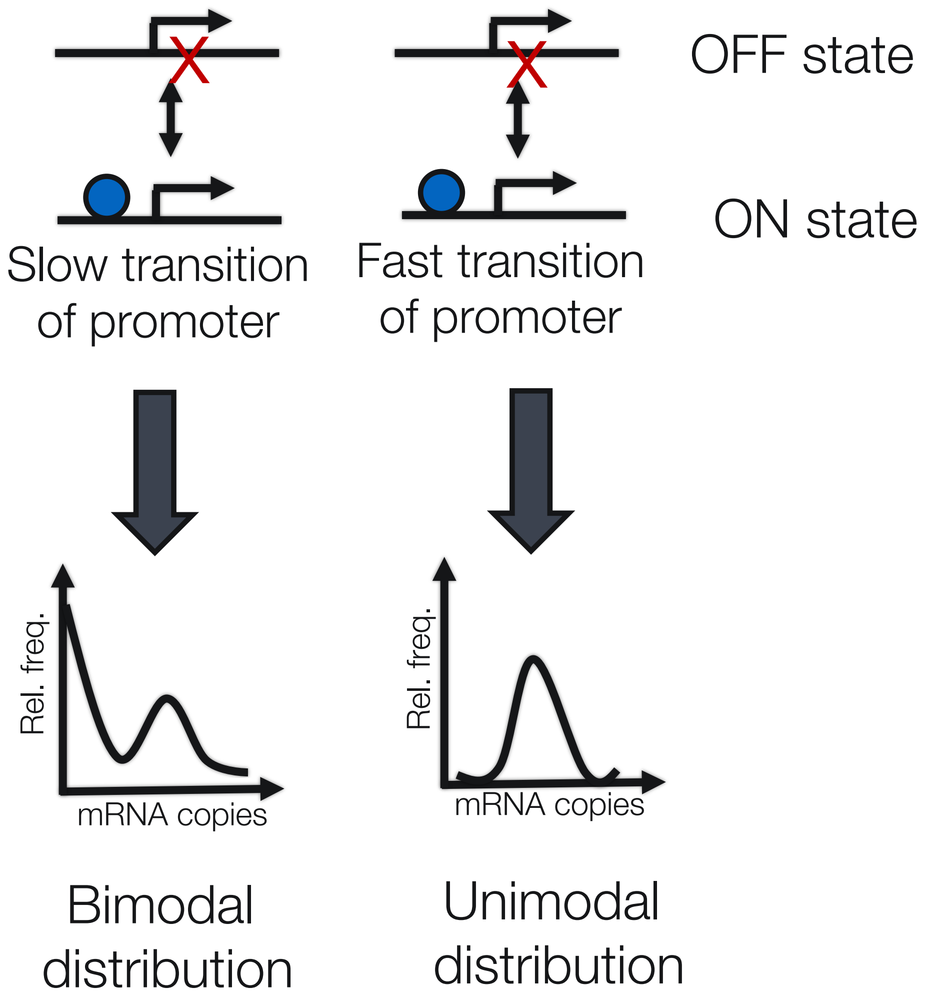
Studying heterogeneity



Lineage tracing study



Stochastic gene expression



Liu S and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges, F1000 Research 2016 (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 2014 (doi: 10.1016/j.cell.2014.02.010)

Introduction to Single Cell RNA Sequencing

- Common applications of single cell RNA sequencing.
- Overview of single cell RNA sequencing platforms.
- Modified scRNA-seq workflows
- Sample preparation and experimental design.
- Effects of sample prep and sample type on analysis.

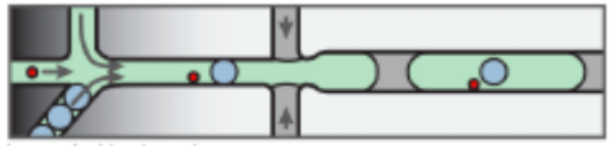
Comparison of Single Cell Methods

CELL-seq
MARS-seq
SMART-seq
SCRB-seq

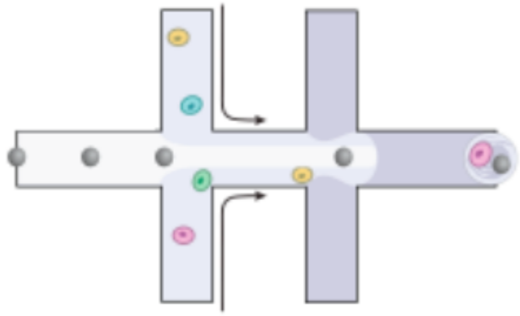
Chromium (10x)

Seq-Well

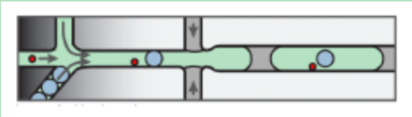
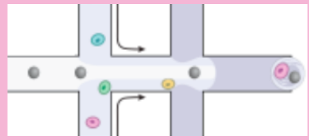
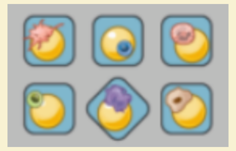
inDrops



Drop-seq

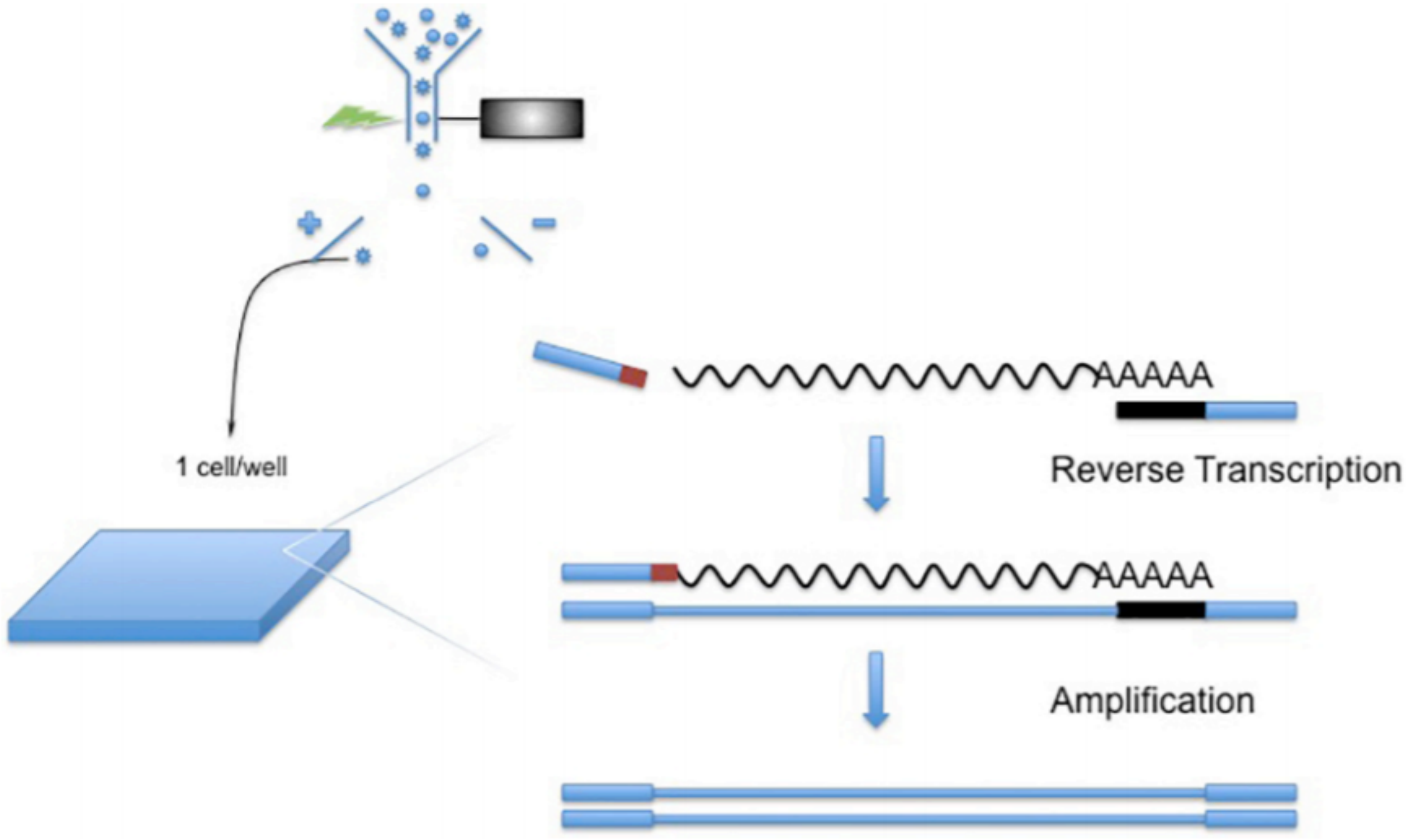


Comparison of Single Cell Methods

	inDrops	10x Genomics	Drop-seq	Seq-well (Honeycomb)	SMART-seq
Cell capture efficiency	~70-80%	~50-70%	~10%	~80%	~80%
Time to capture 10k cells	~30min	10min	1-2 hours	5-10min	--
Encapsulation type	Droplet 	Droplet 	Droplet 	Nanolitre well 	Plate-based 
Library prep	CEL-seq Linear amplification by IVT	SMART-seq Exponential PCR based amplification	SMART-seq Exponential PCR based amplification	SMART-seq Exponential PCR based amplification	SMART-seq Exponential PCR based amplification
Commercial	Yes	Yes	--	Yes (Summer 2020)	Yes
Cost (~$ per cell)	~0.06	~0.2	~0.06	~0.15	1
Strengths	• Good cell capture • Cost-effective • Real-time monitoring • Customizable	• Good cell capture • Fast and easy to run • Parallel sample collection • High gene / cell counts	• Cost-effective • Customizable	• Good cell capture • Cost-effective • Real-time monitoring • Customizable	• Good cell capture • Good mRNA capture • Full-length transcript • No UMI
Weaknesses	Difficult to run	Expensive	Difficult to run & low cell capture efficiency	Available Soon	Expensive

C. Ziegenhain et al., Comparative Analysis of Single-Cell RNA Sequencing Methods, Molecular Cell 2017 (doi: 10.1016/j.molcel.2017.01.023)

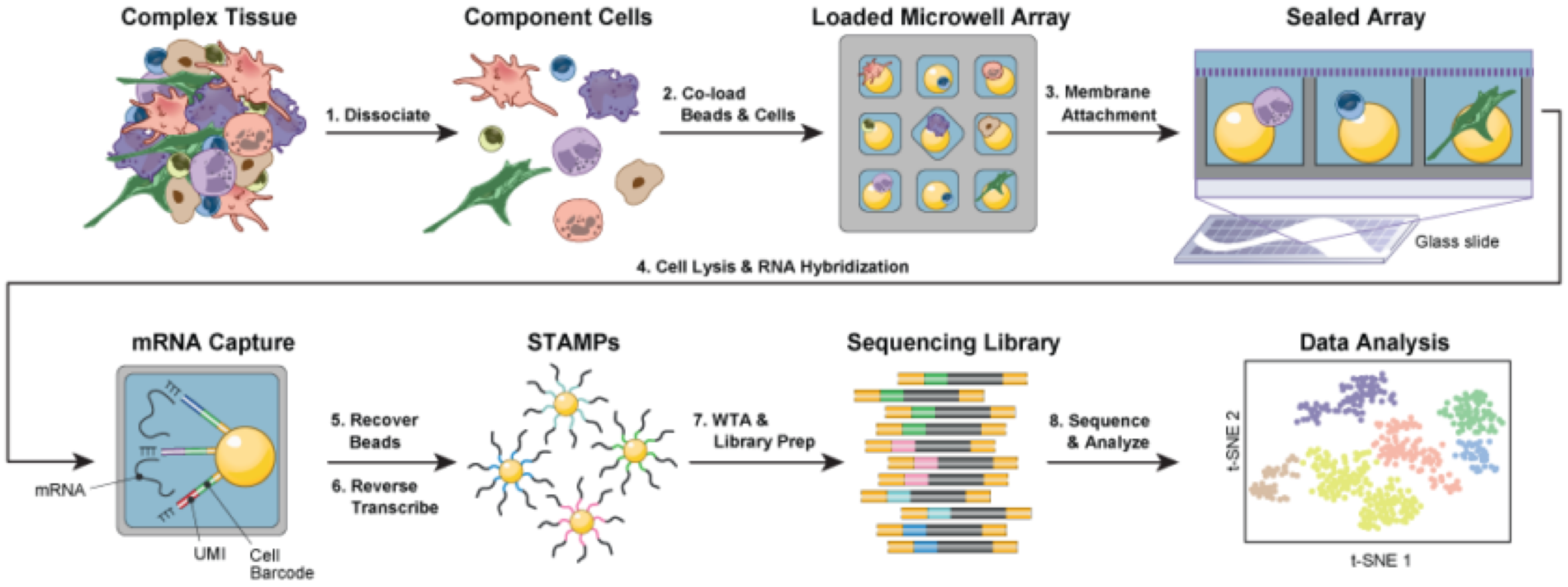
Full Length Transcripts: SMART-seq (v3)

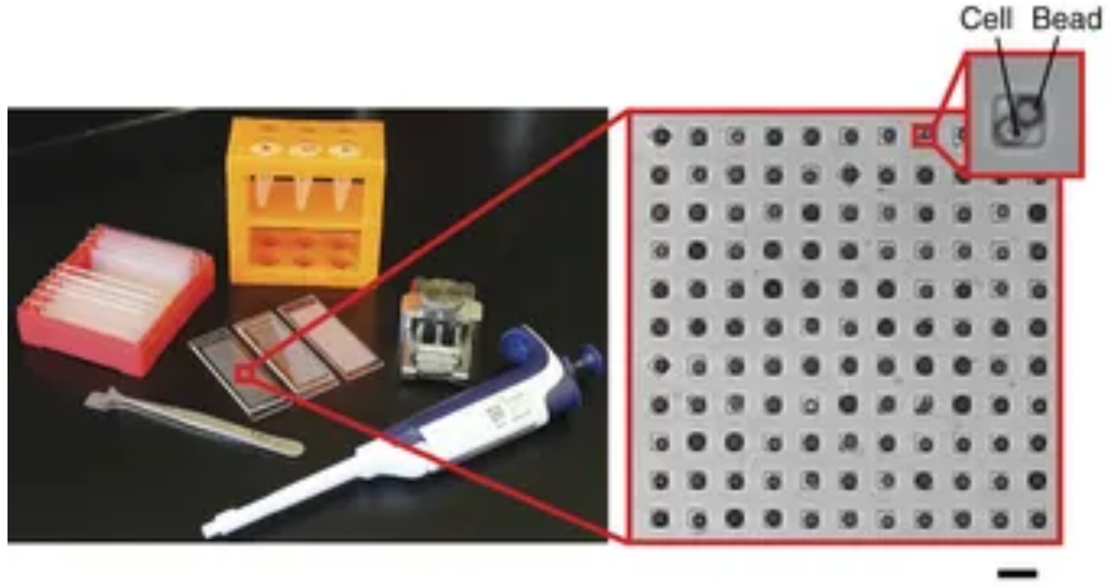


- Sort cells of interest into single well.
- Only single cell method that gives full transcript information.
- Currently best option for low cell number samples. (100’s – 1,000’s)

H Lim et al, Profiling Individual Human Embryonic Stem Cells by Quantitative RT-PCR. J. Vis. Exp. (87), e51408, 2014 (doi:10.3791/51408)
M Hagemann-Jensen et al, Single-cell RNA counting at allele- and isoform-resolution using Smart-seq3 bioRxiv 2019 (doi: https://doi.org/10.1101/817924)

Seq-Well: Honeycomb Biotechnologies





- Up to 1ml of sample is loaded into nanowells (100’s – 1,000’s of cells)
- Samples settle by gravity into wells.
- Uses Drop-seq STAMP beads for barcoding.

TM Gierahn et al, Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Nat Methods. 2017 Apr;14(4):395-398. doi: 10.1038/nmeth.4179

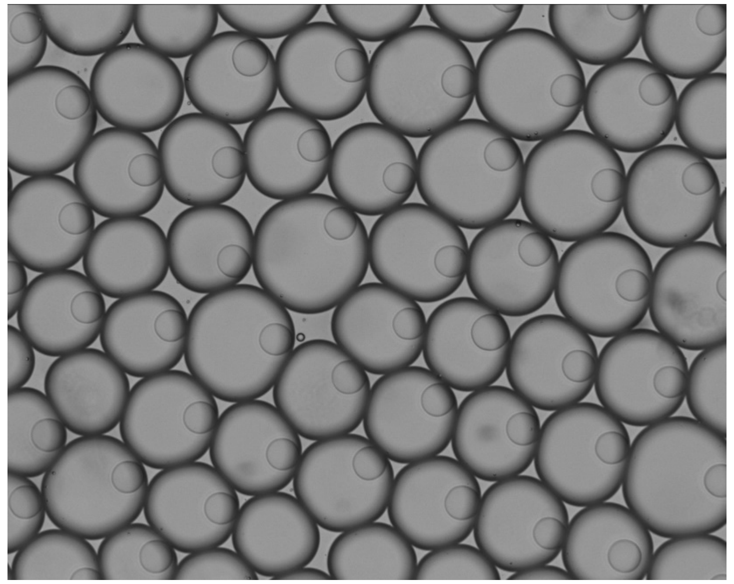
Seq-Well: Honeycomb Biotechnologies

- Honeycomb devices require no complex equipment or training so samples can be can be collected in almost any setting.
- Gentle stabilization process allows even the most fragile cells to be captured.

https://honeycomb.bio/

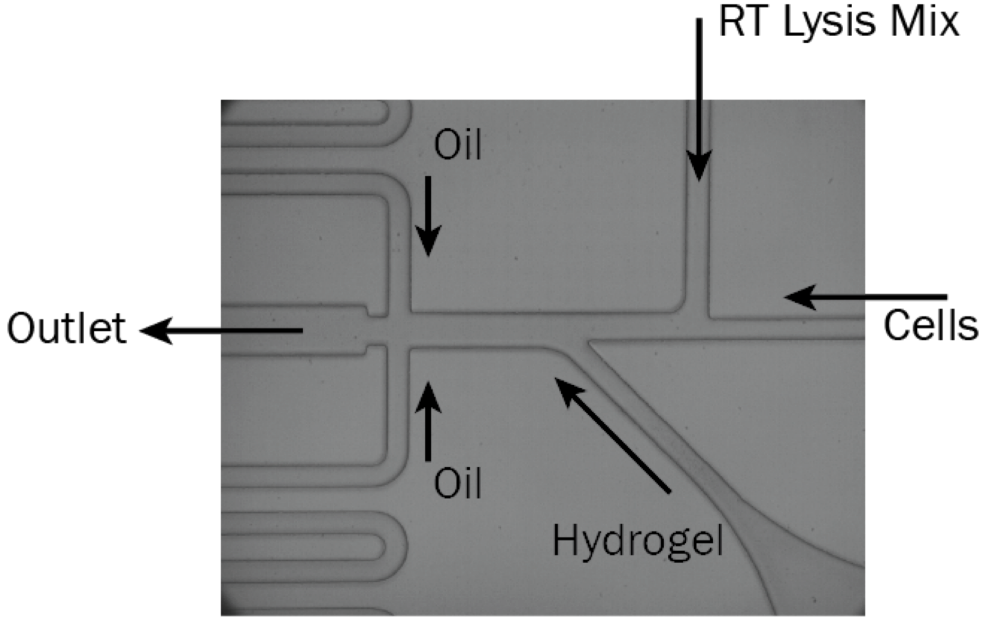
Droplet scRNA-seq

- Droplet methods give single cell information, BUT require high cell numbers to achieve best results. (>10,000 cells)
- Capture 50-80% of the input cells depending on the platform used.
- Extreme 3’-bias in standard data.
- Can not look at splicing / isoforms with standard methods.



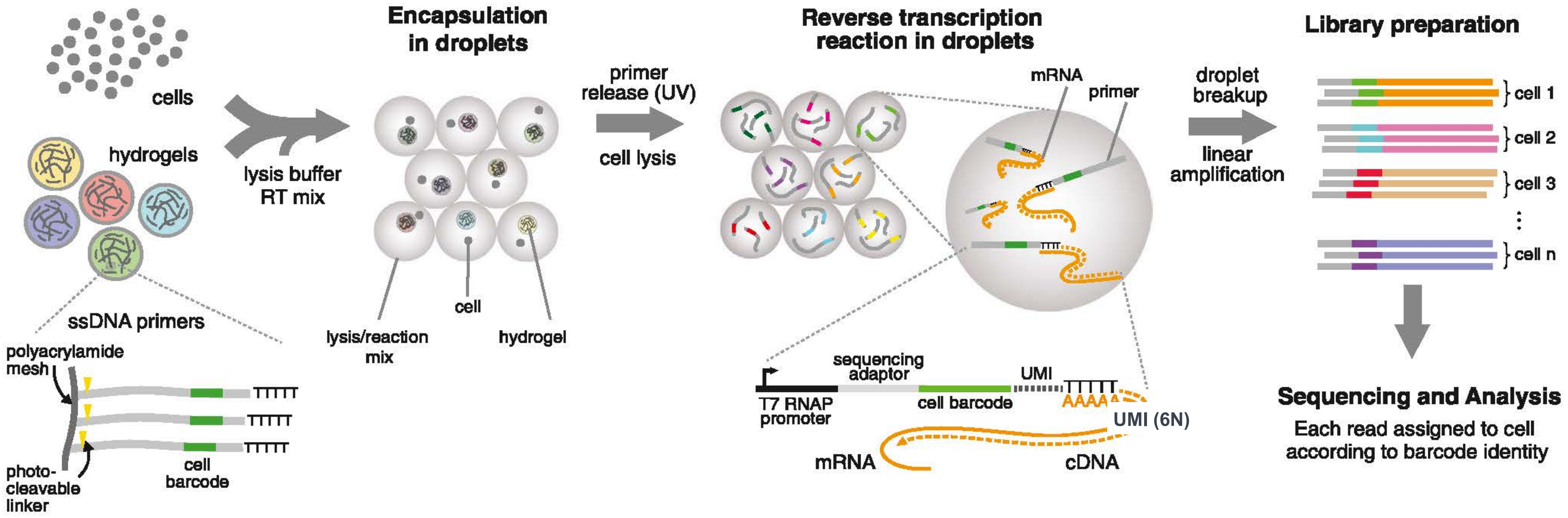
inDrops Method Overview

- Single cell suspension injected at density of ~80,000 cells / ml



- Matching the speed of bead injection with the speed of droplet generation it is possible to set conditions in which nearly every droplet would be loaded.

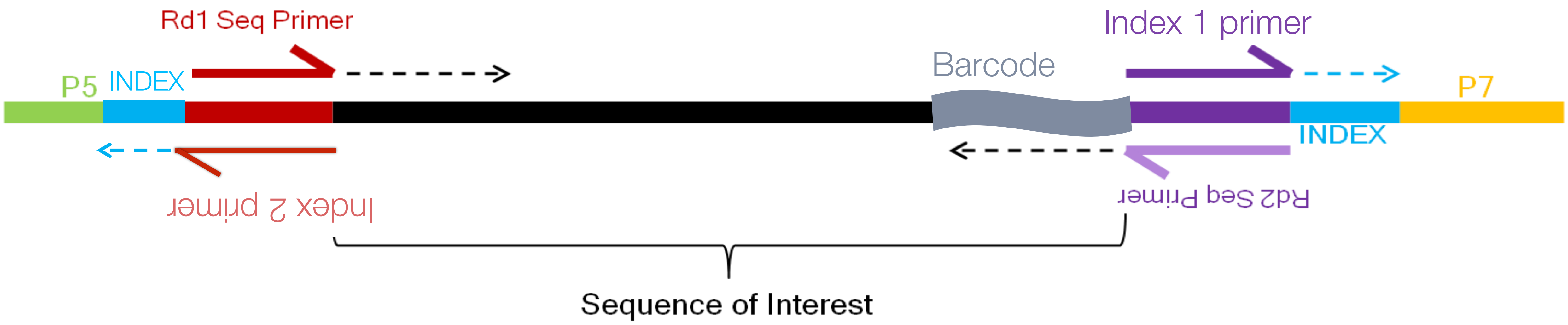
inDrops Method Overview



- Lysis and reverse transcription occurs in the beads
- Samples are frozen after RT as RNA:DNA hybrid in gel
- Library prep is based on CEL-seq method (for now!)
- Moving to template-switching (SMART-seq) style library prep

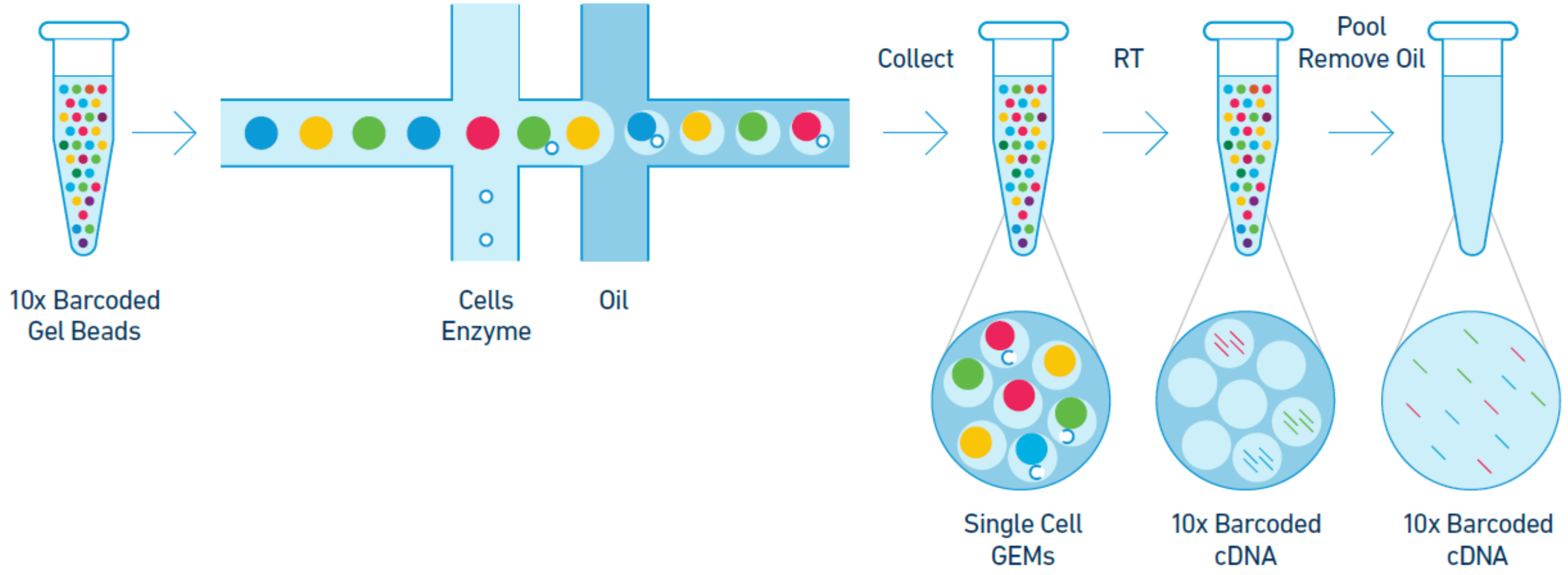
A. M. Klein et al., Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells, Cell 2015 (doi: 10.1016/j.cell.2015.04.044)
R. Zilionis et al., Single-cell barcoding and sequencing using droplet microfluidics, Nature Protocols 2016 (doi: 10.1038/nprot.2016.154)

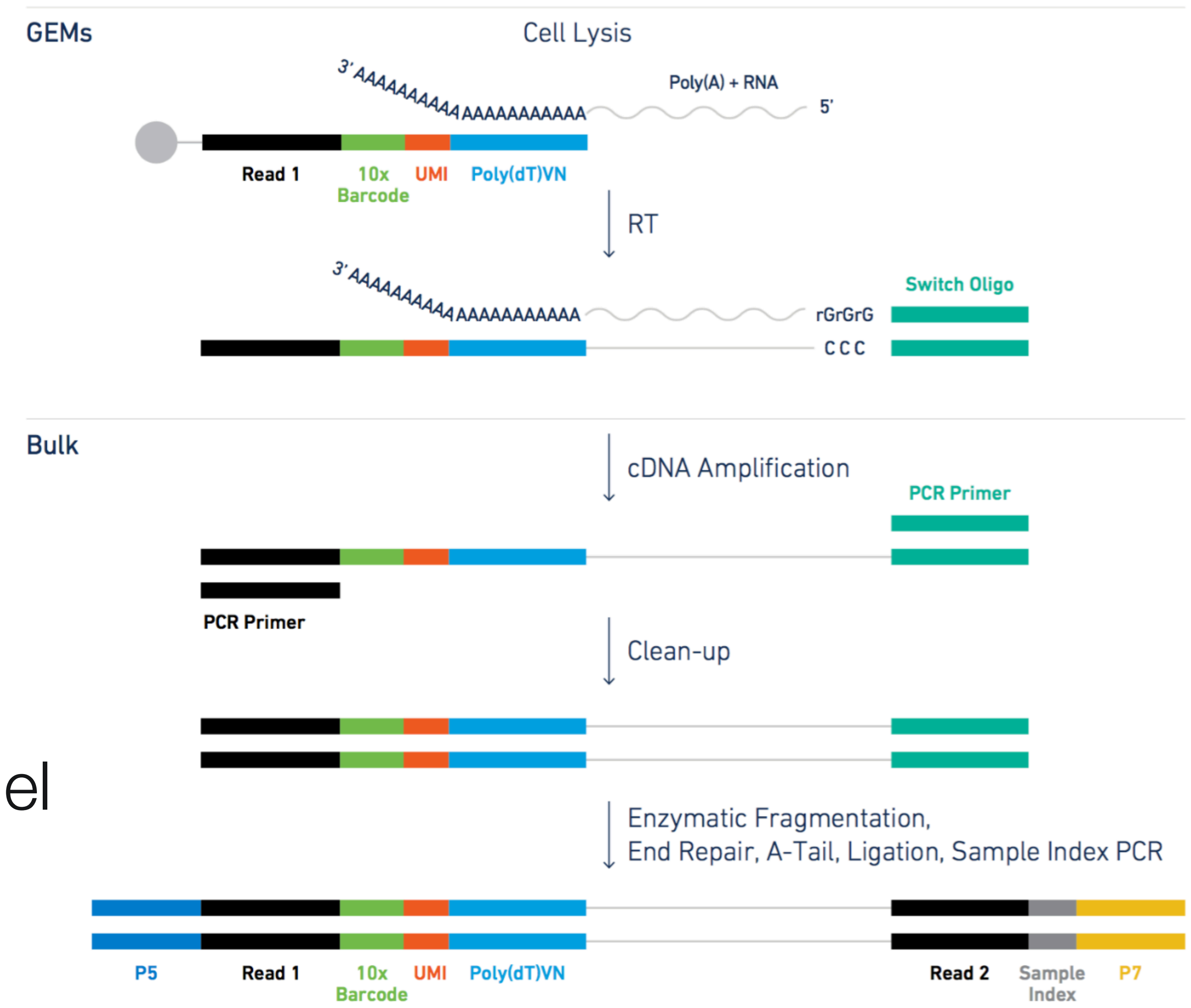
scRNA-seq Library Structure (inDrops)



Sequencing Read	Description	Number of Cycles	Notes
Read1	Insert (Transcript)	61	Can read longer into transcript if desired
I7 index	Single Cell Barcode	8	Reads first half of barcode
I5 index	Library Index	8	Distinguish multiplexed samples
Read 2	Single Cell Barcode & UMI	14	Reading longer will read into PolyA tail.

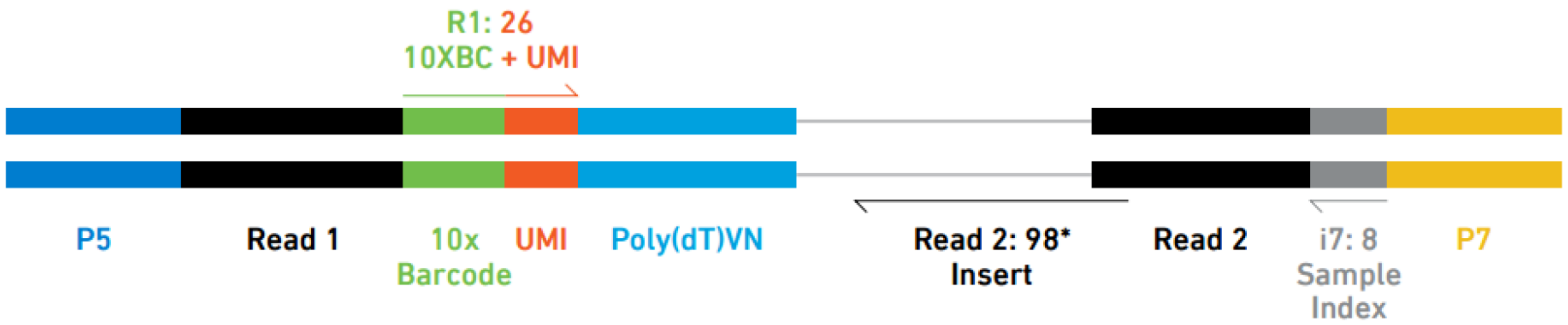
10x Genomics Method Overview





- Lysis and reverse transcription occurs in the beads
- Samples are frozen after RT as RNA:DNA hybrid in gel
- Library prep is similar to SMART-seq method
- GEM = Gel Beads-in-emulsion

10x Genomics Method Overview

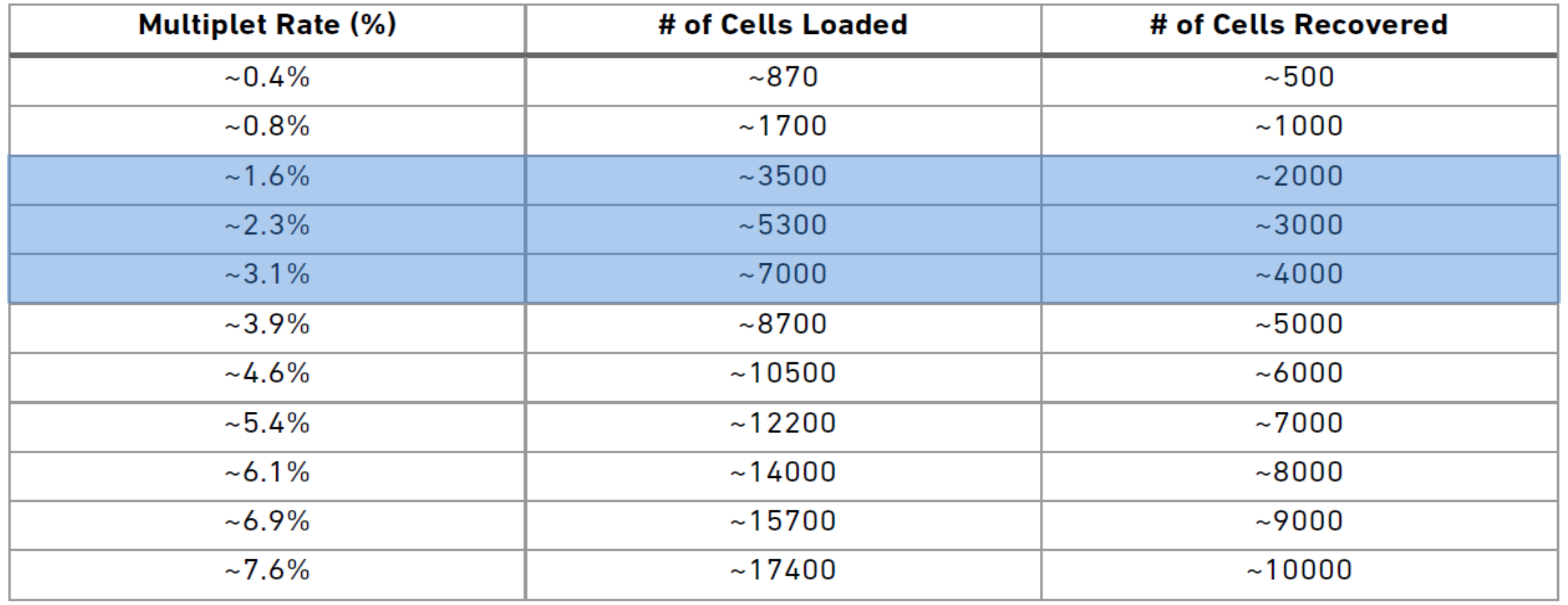


Sequencing Read	Description	Recommended Number of Cycles	Notes
Read 1	10x Barcode Read (Cell) and Randomer Read (UMI)	26 cycles	It cannot be shorter than 26 bp If longer than 26 bp it will be ignored by Cell Ranger
i7 Index	Sample Index Read	8 cycles	If longer than 8 bp it will need to be trimmed during base calling
i5 Index	N/A	0 cycles	N/A
Read 2	Insert Read (Transcript)	98 cycles	Can be adjusted*

**User controlled trade-off between read length and mapping rate*

Doublets / Cell Density

- Rate of doublets depends on the cell density and the flow rate used for encapsulation.

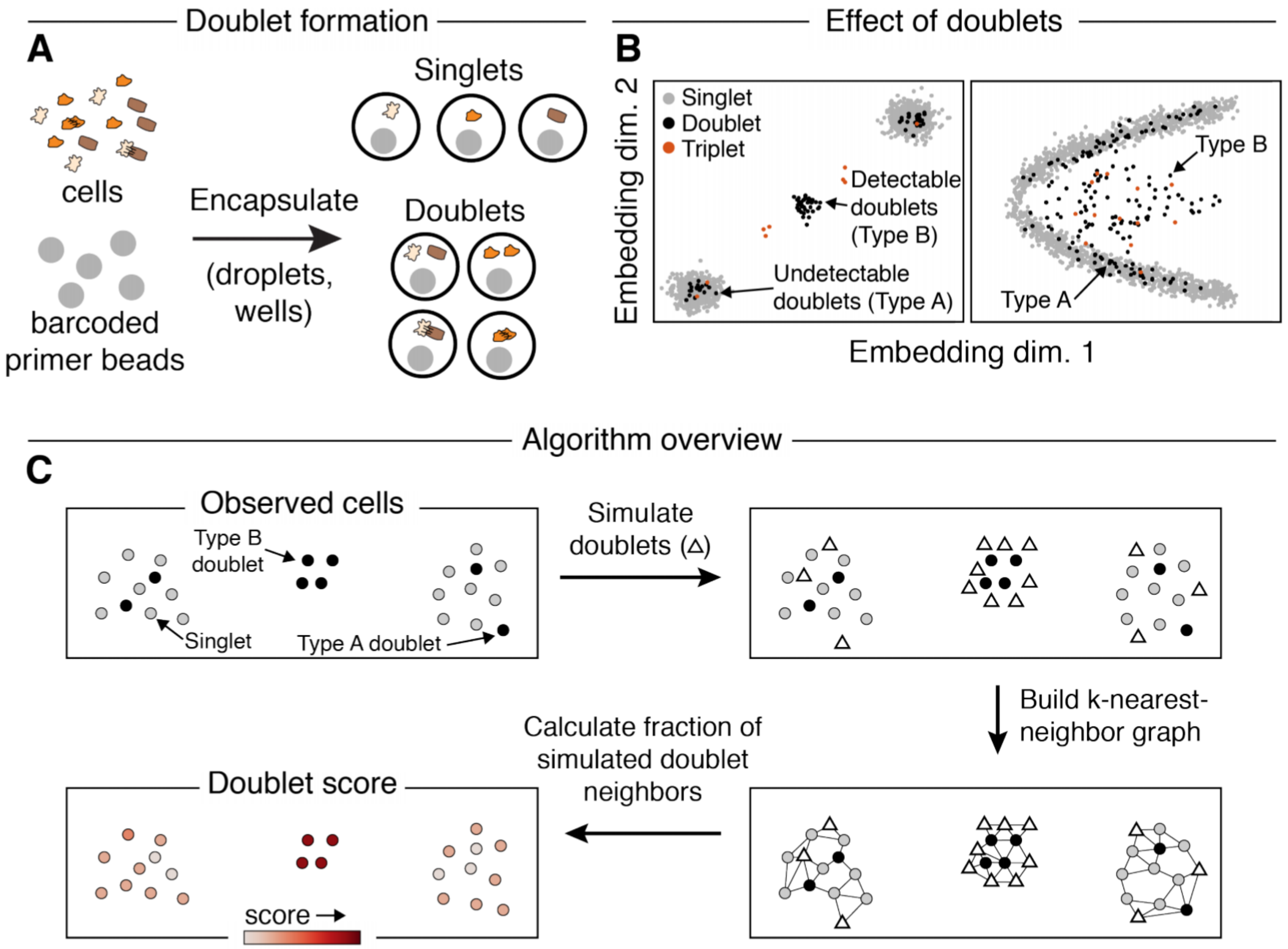


Multiplet Rate (%)	# of Cells Loaded	# of Cells Recovered
~0.4%	~870	~500
~0.8%	~1700	~1000
~1.6%	~3500	~2000
~2.3%	~5300	~3000
~3.1%	~7000	~4000
~3.9%	~8700	~5000
~4.6%	~10500	~6000
~5.4%	~12200	~7000
~6.1%	~14000	~8000
~6.9%	~15700	~9000
~7.6%	~17400	~10000

- “Sweet spot” for loading a 10x is recovery of 5000-6000 cells.

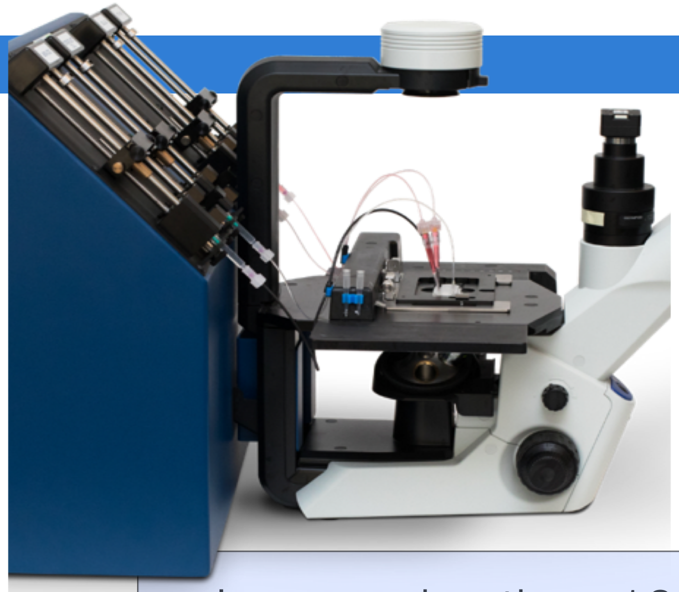
https://support.10xgenomics.com/single-cell-gene-expression/sample-prep/doc/user-guide-chromium-single-cell-3-reagent-kit-v2-chemistry

Scrublet: Computational Identification of Doublets



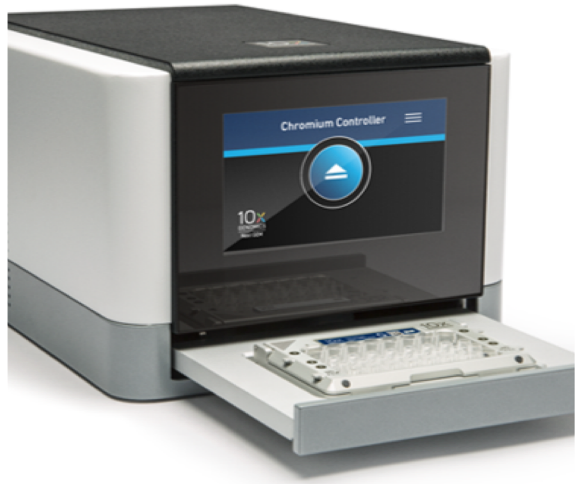
S. Wolock et al. Scrublet: computational identification of cell doublets in single-cell transcriptomic data, bioRxiv 2018 (DOI: 10.1101/357368)

inDrops vs. 10x Genomics



inDrops

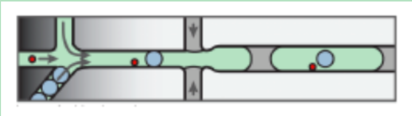
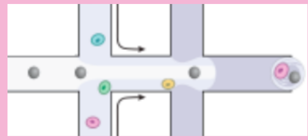
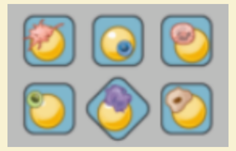
- Lower price than 10x
- Sample cost depends on number of cells and libraries prepped.
- Ability to collect sample backups at reduced cost per sample.
- Samples run 1 at a time for ~20 minutes per 6000 cell sample
- Observe ~1/2 the genes/cell of 10x (old library prep method)
- Customizable



10x Genomics

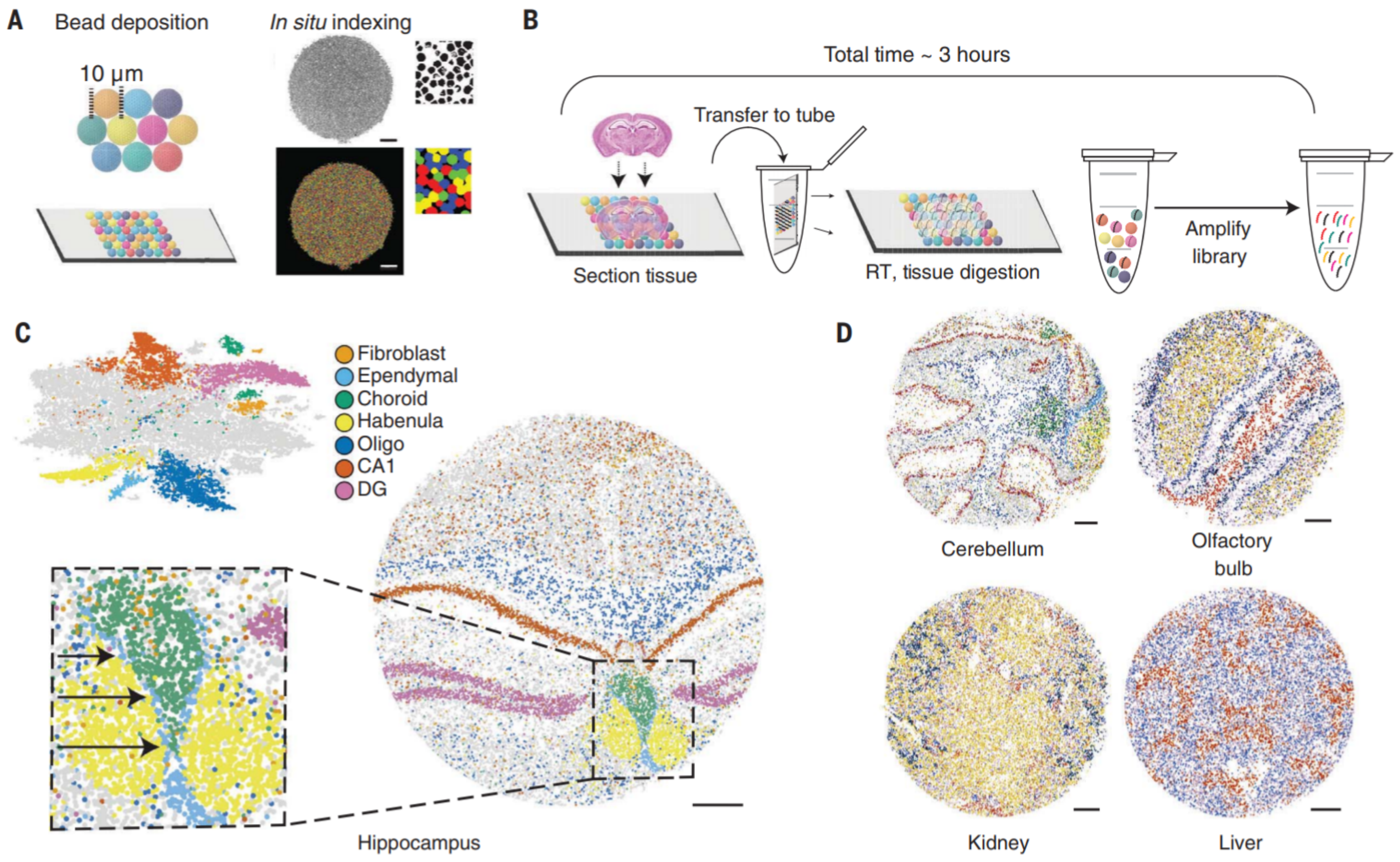
- Higher price, reduced with hashing.
- Sample cost is per sample, number of cells not part of cost.
- Backup samples come at full sample cost.
- Up to 8 samples run in parallel in 10 minutes
- Observe 2x genes/cell than inDrops
- Out-of-the-Box

Comparison of Single Cell Methods

	inDrops	10x Genomics	Drop-seq	Seq-well (Honeycomb)	SMART-seq
Cell capture efficiency	~70-80%	~50-70%	~10%	~80%	~80%
Time to capture 10k cells	~30min	10min	1-2 hours	5-10min	--
Encapsulation type	Droplet 	Droplet 	Droplet 	Nanolitre well 	Plate-based 
Library prep	CEL-seq Linear amplification by IVT	SMART-seq Exponential PCR based amplification	SMART-seq Exponential PCR based amplification	SMART-seq Exponential PCR based amplification	SMART-seq Exponential PCR based amplification
Commercial	Yes	Yes	--	Yes (Summer 2020)	Yes
Cost (~$ per cell)	~0.06	~0.2	~0.06	~0.15	1
Strengths	• Good cell capture • Cost-effective • Real-time monitoring • Customizable	• Good cell capture • Fast and easy to run • Parallel sample collection • High gene / cell counts	• Cost-effective • Customizable	• Good cell capture • Cost-effective • Real-time monitoring • Customizable	• Good cell capture • Good mRNA capture • Full-length transcript • No UMI
Weaknesses	Difficult to run	Expensive	Difficult to run & low cell capture efficiency	Available Soon	Expensive

C. Ziegenhain et al., Comparative Analysis of Single-Cell RNA Sequencing Methods, Molecular Cell 2017 (doi: 10.1016/j.molcel.2017.01.023)

On the Horizon: Spatial Transcriptomics



- All methods are in their infancy.
- Require fresh-frozen tissue sections.
- Kharchenko lab working on setting up Slide-seq (10 micron resolution)
- Slide-seq recovers 100’s genes/cell
- 10x Visium (50-100 micron resolution)

Rodriques et al, Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. Science. 2019 Mar 29;363(6434):1463-1467.

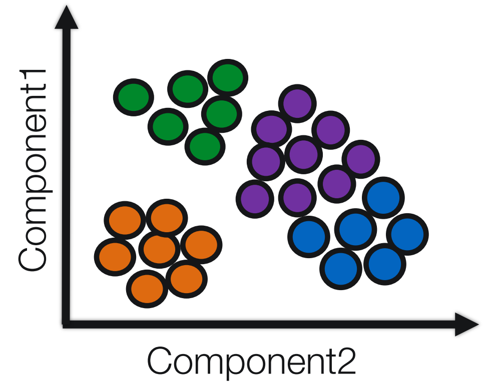
Introduction to Single Cell RNA Sequencing

- Common applications of single cell RNA sequencing.
- Overview of single cell RNA sequencing platforms.
- Modified scRNA-seq workflows
- Sample preparation and experimental design.
- Effects of sample prep and sample type on analysis.

Transcript Specific Library Prep

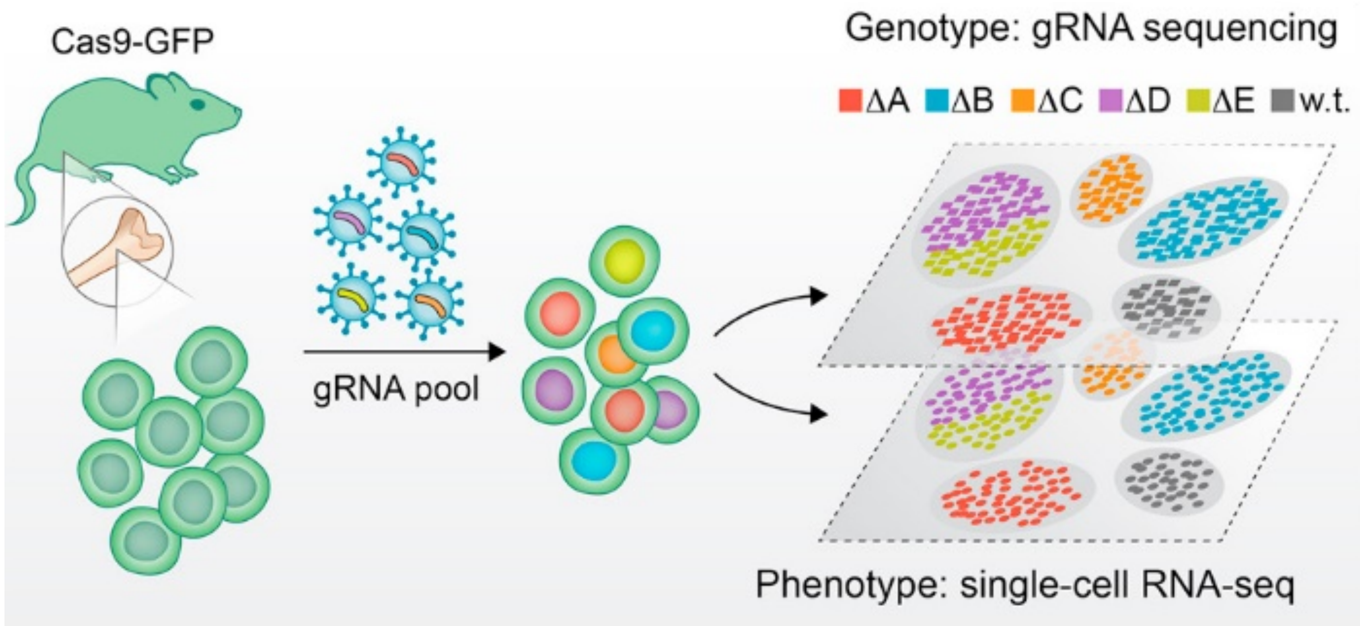
- Identify cells in pooled CRISPR screens
- Identify barcoded cells
- V(D)J immune cell profiling
- Identify labeled (GFP, mCherry) cells.
- Enrich for genes of particular interest for your experiment.

GFP labeled population



Transcript Specific Library Prep

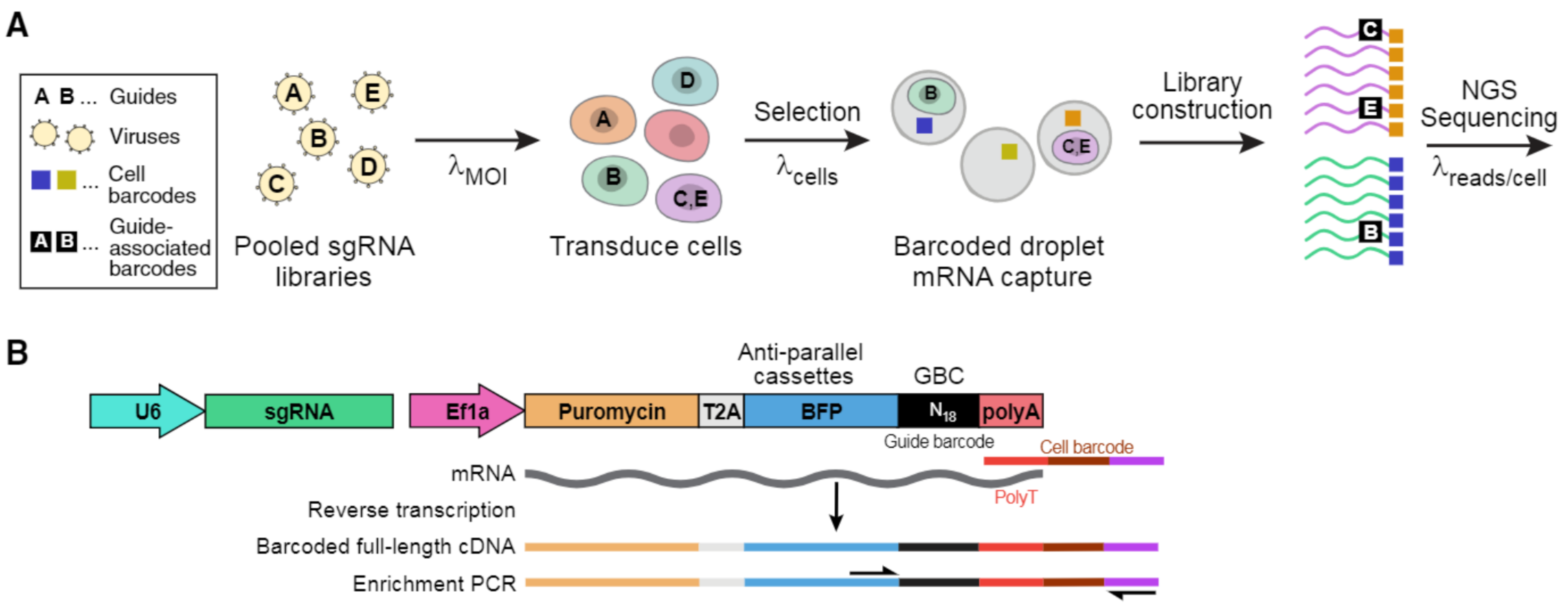
CRISPR pool vector backbone must contain a transcribed poly-adenylated unique guide index (UGI), which can include a fluorescent marker



- scRNA-seq library to phenotype cellular transcriptome
- gRNA-targeted library to ensure proper cell identification

D.A. Jaitin et al., Dissecting Immune Circuits by Linking CRISPR Pooled Screens with Single-Cell RNA-Seq, Cell 2016 (doi: 10.1016/j.cell.2016.11.039)
A. Dixit et al., Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens, Cell 2016 (doi10.1016/j.cell.2016.11.038)

Transcript Specific Library Prep

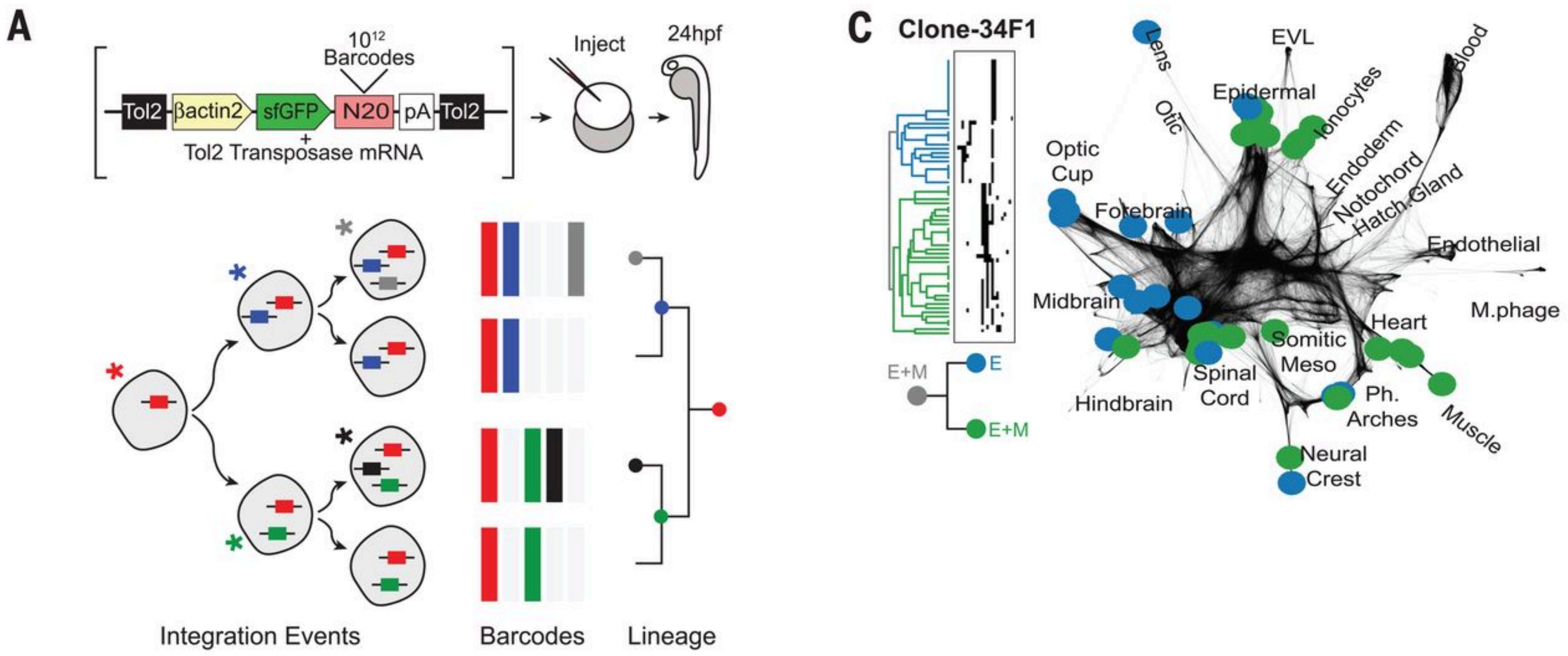


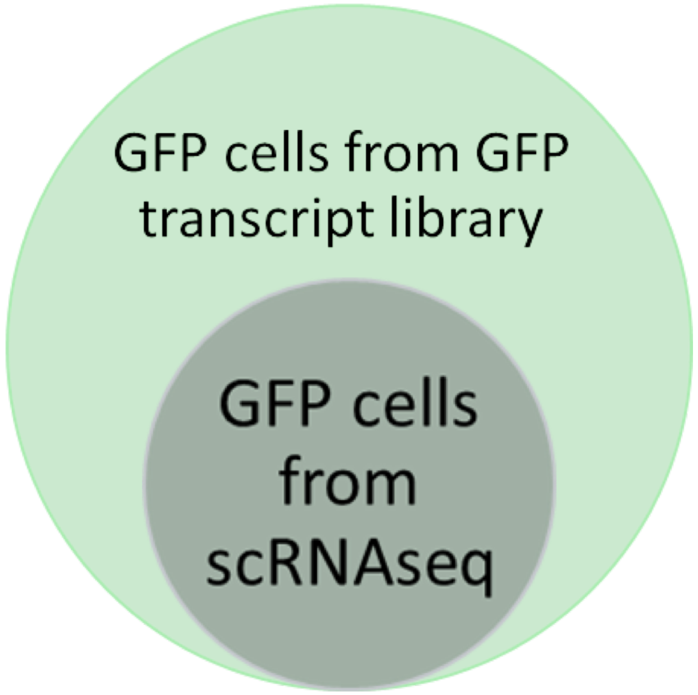
A. Dixit et al., Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens, Cell 2016 (doi10.1016/j.cell.2016.11.038)

Transcript Specific Library Prep

- Make standard library
  - NextSeq/HiSeq/NovaSeq sequencing to identify cell barcodes in sample
- Make transcript specific library with aliquot of partially prepped library
  - MiSeq to identify cells with transcripts of interest or spike back into full transcriptome sample
  - Can also spike the transcript specific library back into full transcriptome library.
- Match barcodes identified in both sequencing runs

Transcript Specific Library Prep





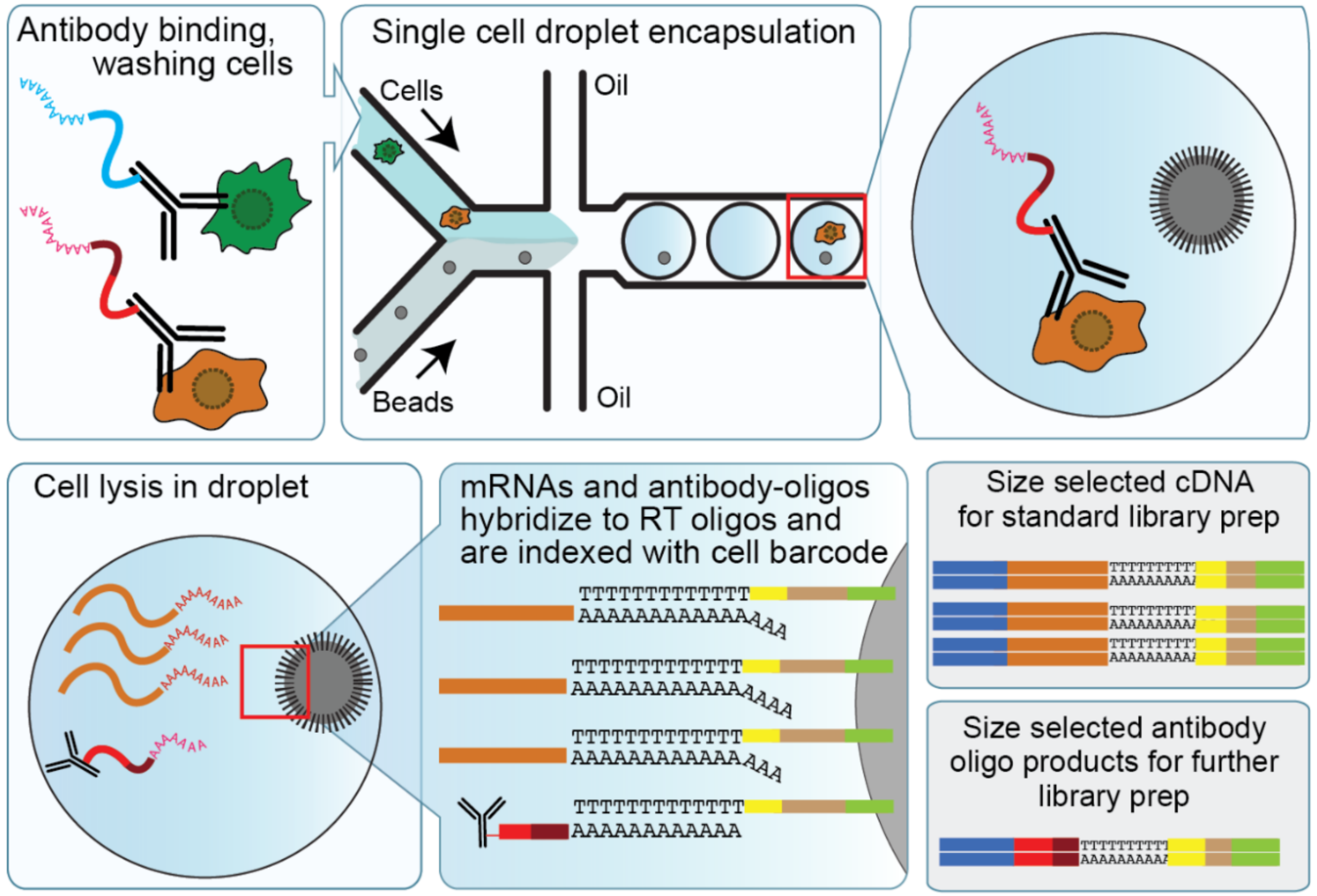
GFP cells from GFP transcript library

GFP cells from scRNAseq

- Transcript of interest often not observed all cells it was expressed in.
  - Specifically amplify desired transcript to identify more cells of interest.
- Recover double the number of GFP cells with the addition of a transcript specific library.

Wagner DE et al., Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. Science. 2018 Jun 1;360(6392):981–987 (doi: 10.1126/science.aar4362)

CITE-seq / Cell Hashing

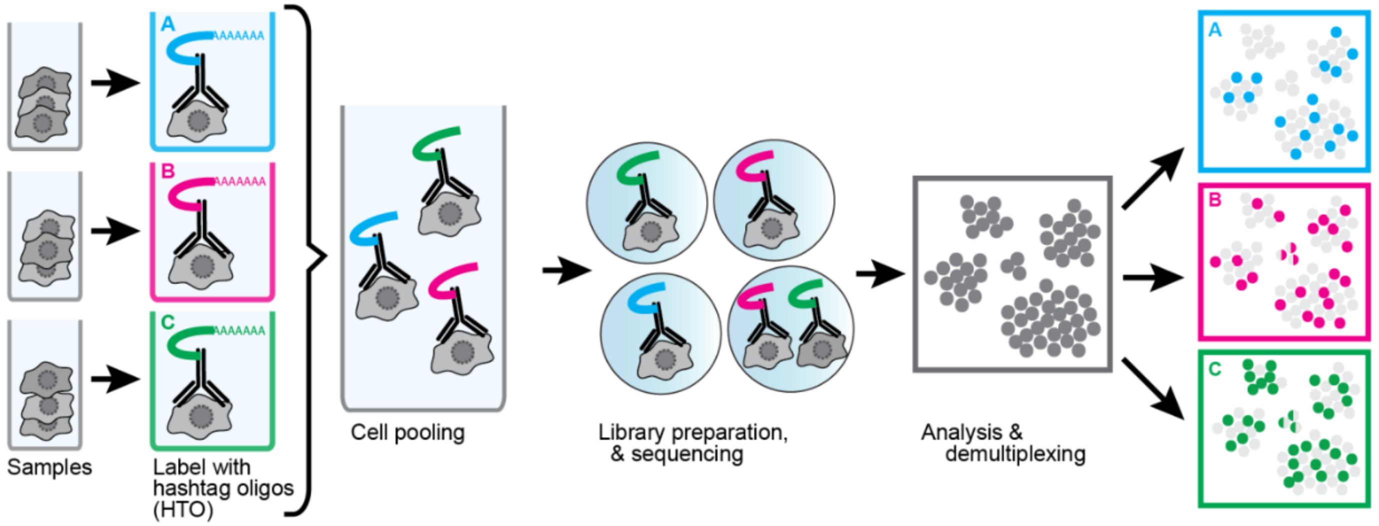


- Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq)
- CITE-seq uses DNA-barcoded antibodies to convert detection of proteins into a quantitative, sequenceable readout.

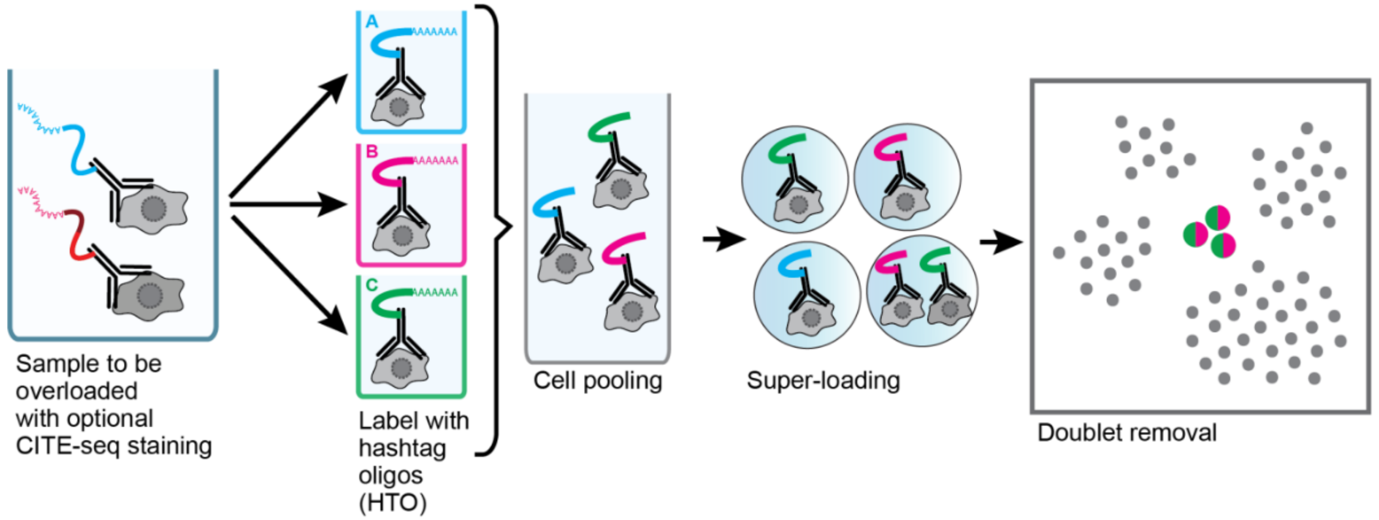
https://cite-seq.com/

Cell Hashing / CITE-seq

Sample multiplexing schematic:



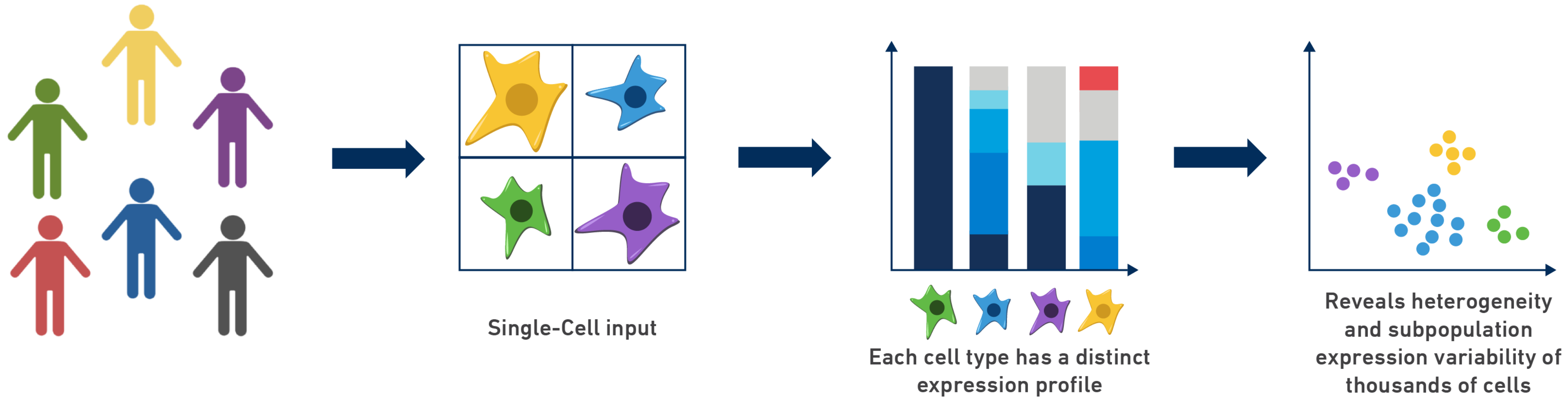
Sample super-loading schematic:



- Cell Hashing is the same idea as CITE-seq just using ubiquitously expressed surface proteins.
- Allows for multiplexing samples into a single encapsulation.
- By sequencing tags alongside the cellular transcriptome, we can assign each cell to its sample of origin, and robustly identify doublets originating from multiple samples.

https://cite-seq.com/

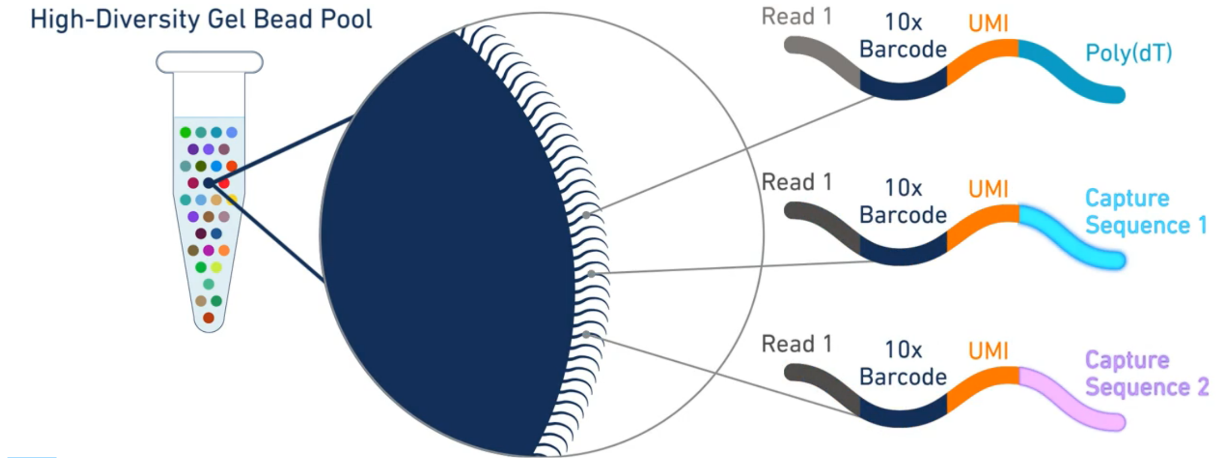
Label-Free Multiplexing of Patient Samples

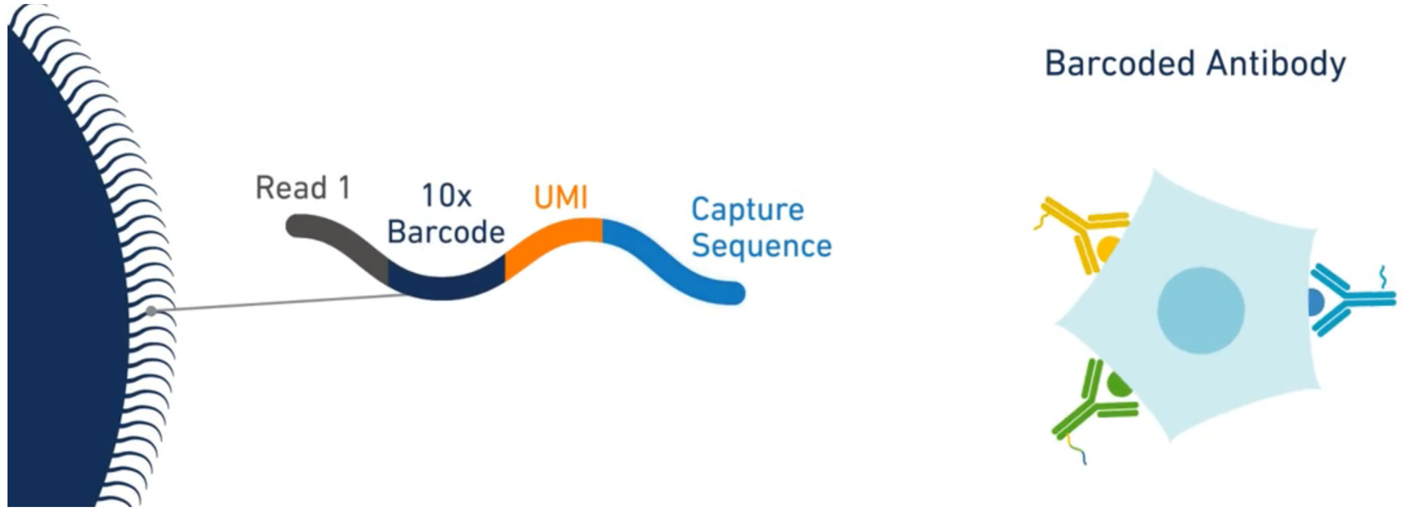


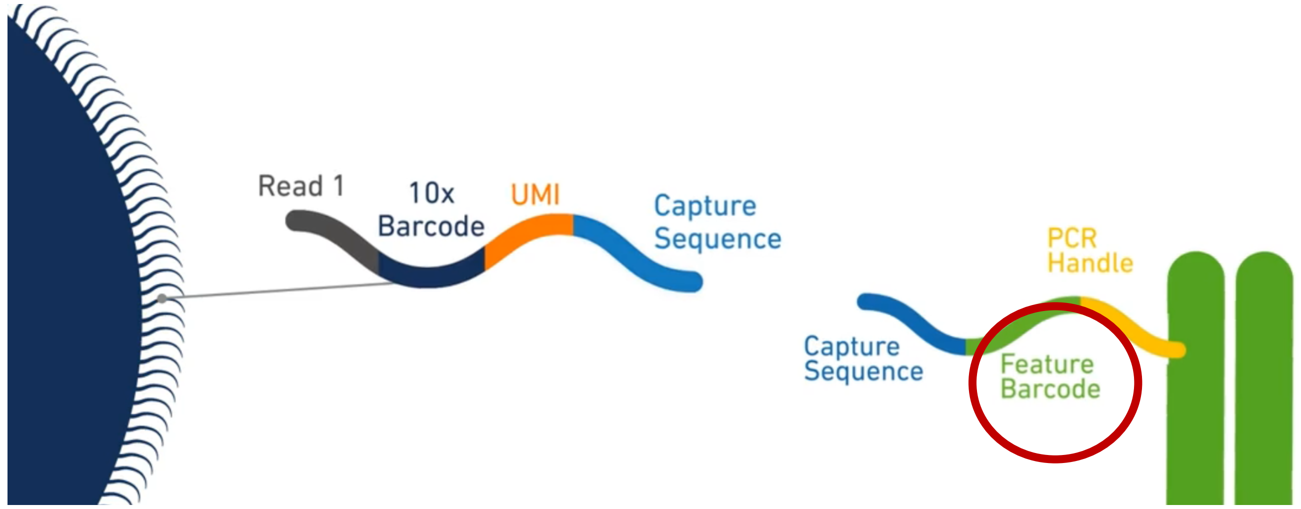
- Genetic differences between patient samples are inferred directly from scRNA-seq data to demultiplex samples.
- Will not be able to link patient data back unless have data on patient SNPs.
- Allows for super-loading and doublet identification.

Xu J. et al., Genotype-free demultiplexing of pooled single-cell RNA-seq. bioRxiv 2019 Mar. 7 (http://dx.doi.org/10.1101/570614)

10x Capture Sequence / Feature Barcode



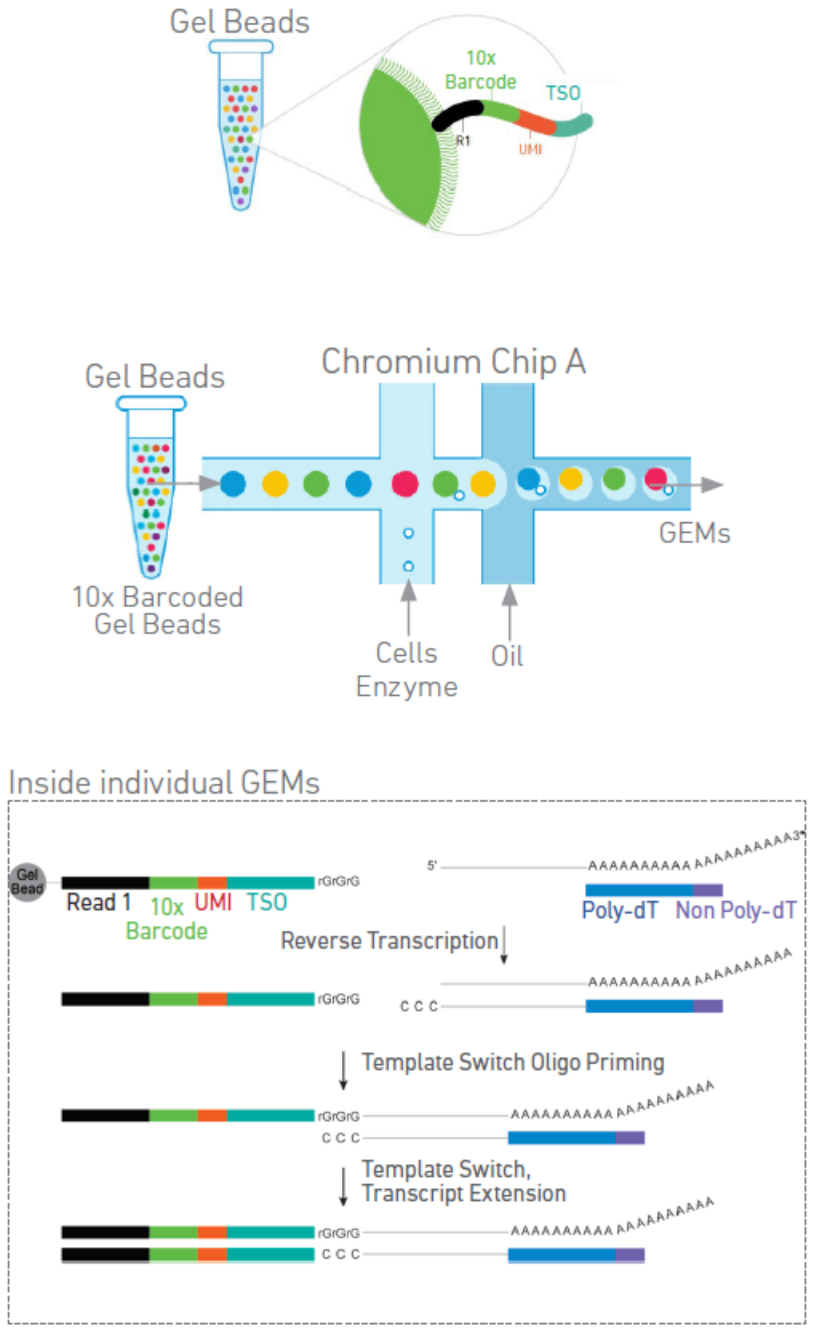




- V3 10x kits offer specific capture sequences you can engineer into your system.
- CITE-seq, CRISPR pools, etc.

https://www.10xgenomics.com/solutions/single-cell/

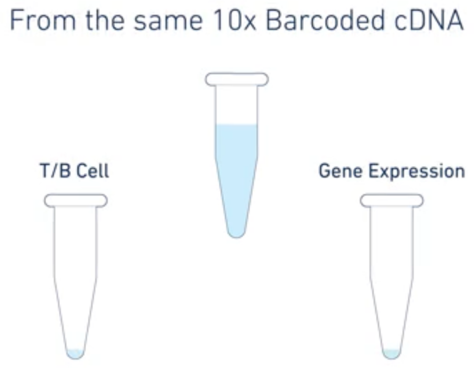
10x V(D)J Immune Profiling & 5’ gene expression

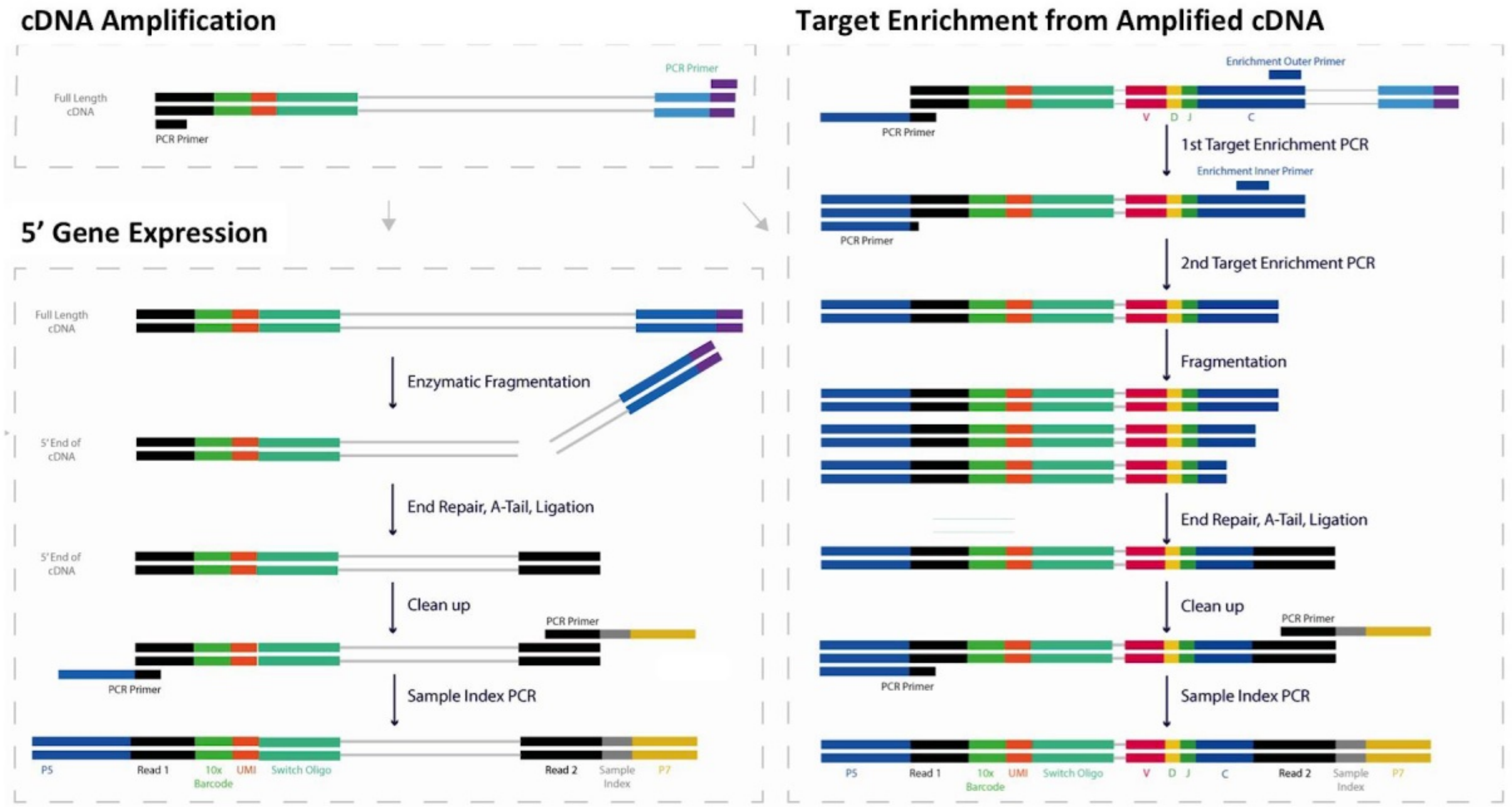


- 3’ gene expression assay the polydT sequence attached to gel bead.
- 5' assay the polydT is supplied as an RT primer.
- 5’ VDJ kit: profiles full length (5’ UTR to constant region) paired T-cell receptor (TCR), or B-cell immunoglobulin (Ig) transcripts.
  - If both T and B cells are expected to be present in the cell population, TCR and Ig transcripts can be enriched in separate reactions from the same amplified cDNA material.

https://kb.10xgenomics.com/hc/en-us/articles/360000939852-What-is-the-difference-between-Single-Cell-3-and-5-Gene-Expression-libraries-

10x V(D)J Immune Profiling

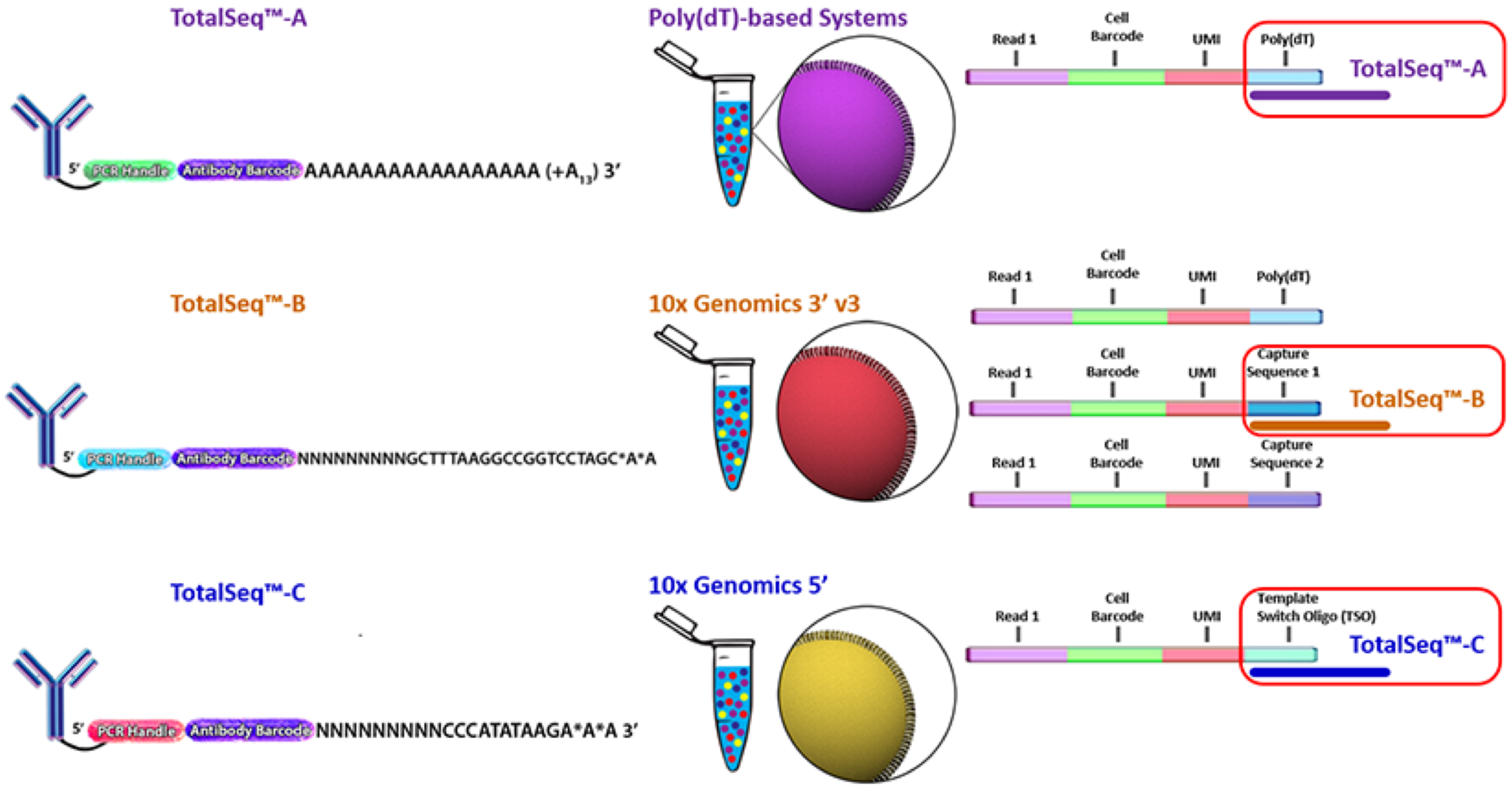




https://www.10xgenomics.com/solutions/vdj/

TotalSeq

- TotalSeq – A: captures using a polyA Tail; 10x 3’ transcriptome V2 or V3
- TotalSeq – B: captures using sequence 1; 10x 3’ transcriptome V3
- TotalSeq – C: captures using sequence TSO; 10x 5’ transcriptome V3



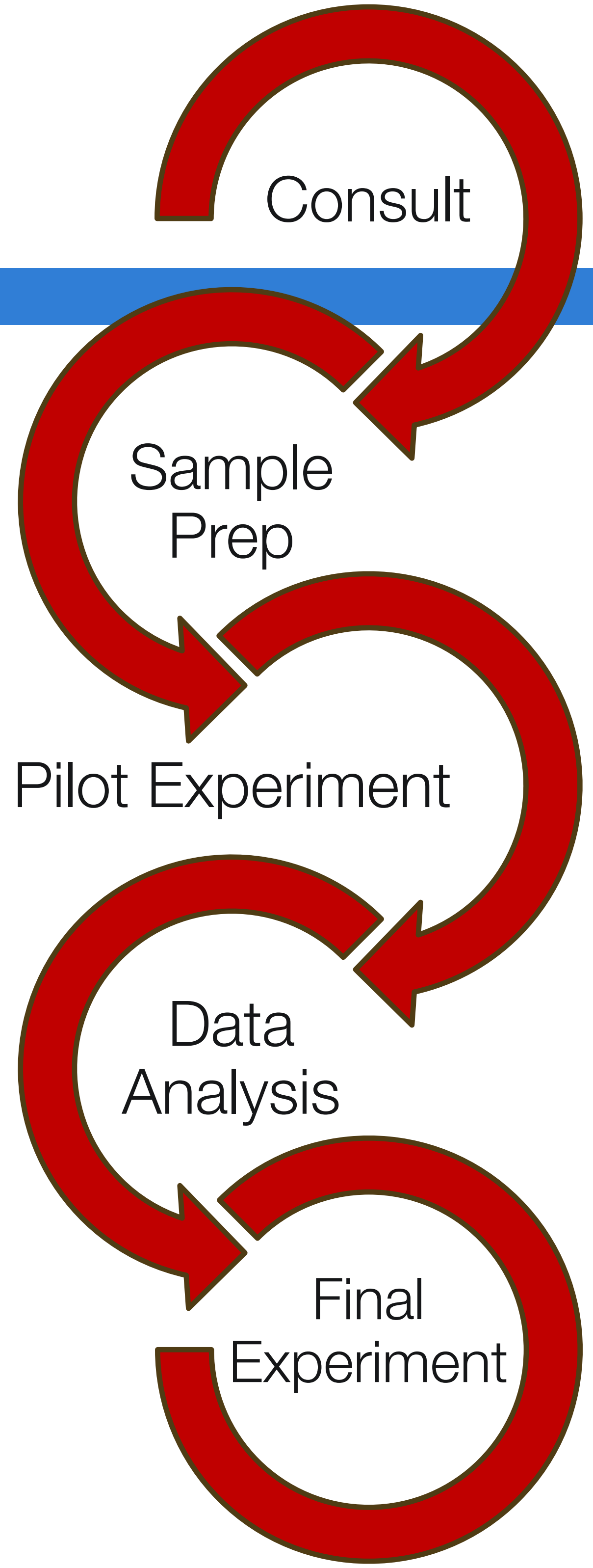
https://kb.10xgenomics.com/hc/en-us/articles/360019665352-What-is-the-difference-between-TotalSeq-A-B-and-C-
https://www.biolegend.com/en-us/totalseq

Introduction to Single Cell RNA Sequencing

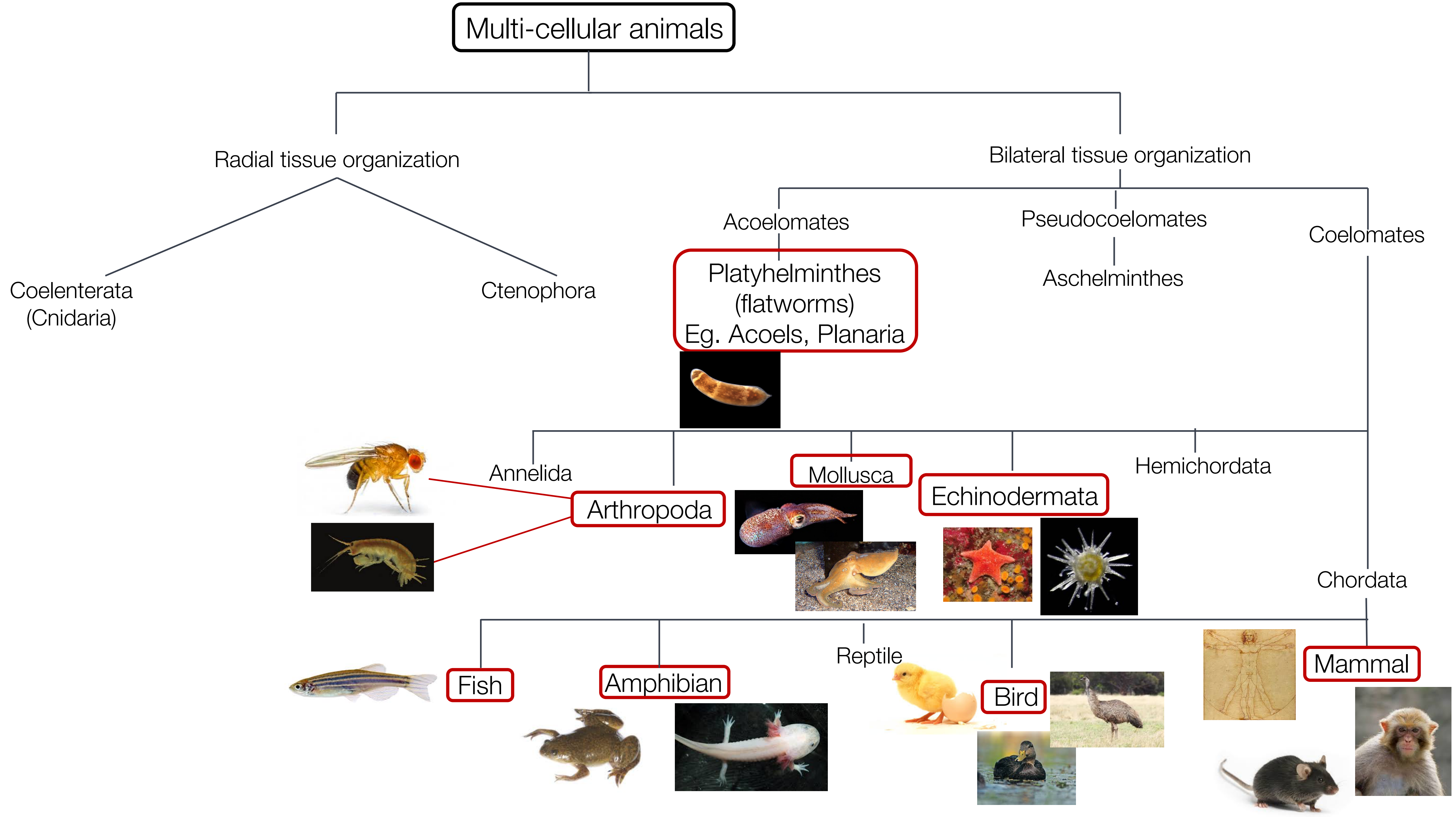
- Common applications of single cell RNA sequencing.
- Overview of single cell RNA sequencing platforms.
- Modified scRNA-seq workflows
- Sample preparation and experimental design.
- Effects of sample prep and sample type on analysis.

Single Cell Core Workflow

- Good sample prep is the key to success.
- A well planned pilot experiment is essential for evaluating sample preparation and for understanding the required number of cells.
- Do not rush to the final experiment.



Single Cell Core Sample Repertoire



Multi-cellular animals

Radial tissue organization

Coelenterata (Cnidaria)

Ctenophora

Bilateral tissue organization

Acoelomates

Platyhelminthes (flatworms) Eg. Acoels, Planaria

Pseudocoelomates

Aschelminthes

Coelomates

Annelida

Arthropoda

Mollusca

Echinodermata

Hemichordata

Chordata

Fish

Amphibian

Reptile

Bird

Mammal

Single Cell Core Sample Repertoire

Primary cell/ tissue:

- Whole embryo
- Blood (immune cells)
- Bone marrow (hematopoietic stem cells)
- Embryonic stem cells
- Gut, Lung, Stomach, Colon (epithelium)
- Adipose tissue (non-adipocytes)
- Hepatocytes
- Brain and spinal cord (neurons, microglia, astrocytes)
- Thymus
- Various patient samples

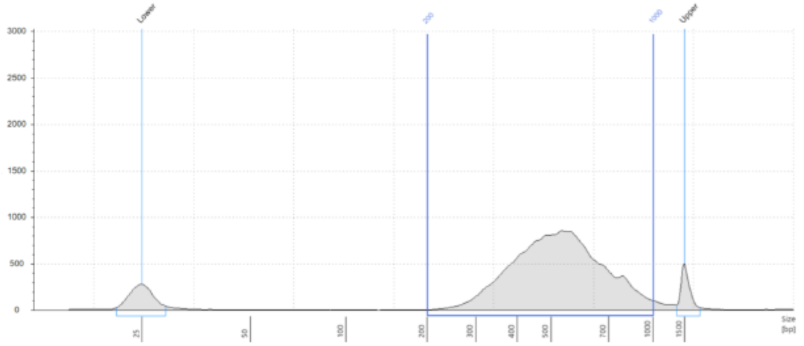
Cells/ tissues grown in vitro:

- Differentiated cells from iPSCs
- Organoids

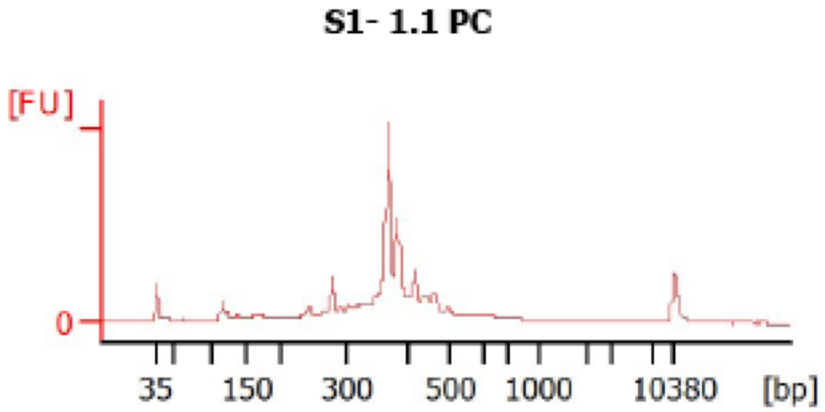
Single nuclear samples:

- Brain and Kidney

Standard Library



Fly / Arthropod Library

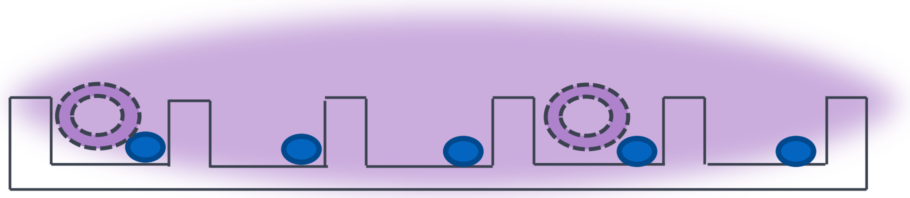


Key to Success: Sample Preparation

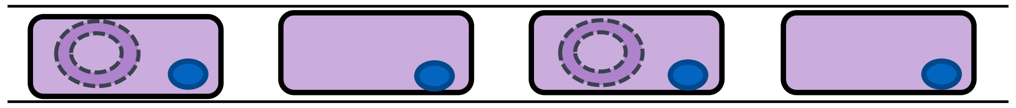
- High cell viability (>90-95% preferred)
- Minimal free-floating RNA
- Single cell suspension
- Dissociation protocol is cell type dependent
- Primary samples are much more difficult
- Cryopreservation or Nuc-seq are optional for some samples

Key to Success: Sample Preparation

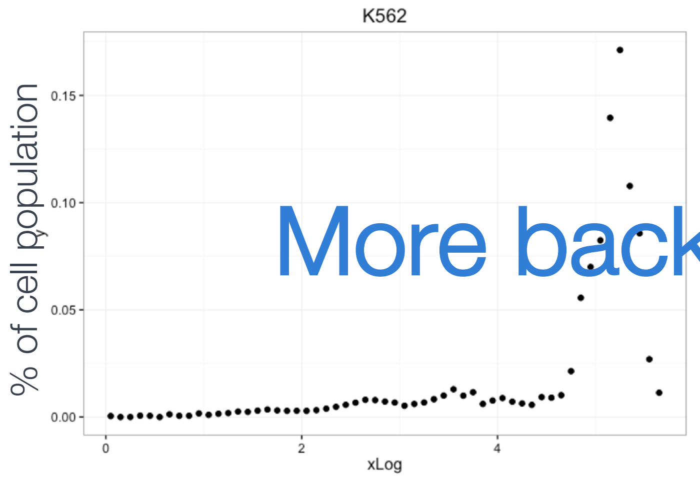
- Good single cell suspension. No clumps and minimal debris.
- More than 90% viability is ideal.
- Cell membrane integrity is required until they are encapsulated.



Nanowell-based barcoding platform

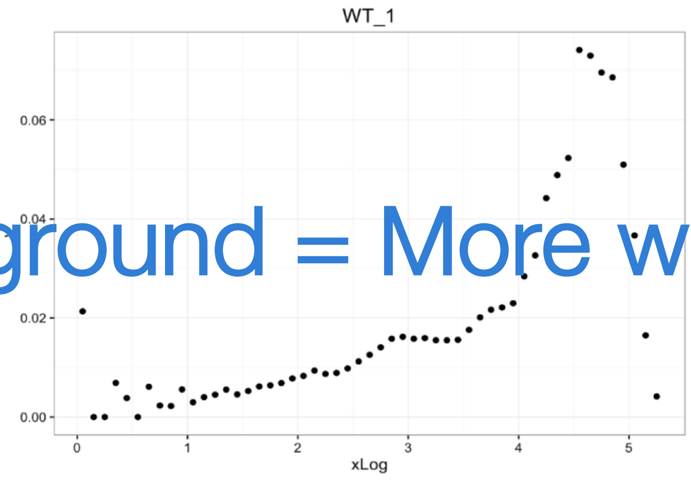


Droplet-based barcoding platform



Reads/ cell barcode

Ideal data



Reads/ cell barcode

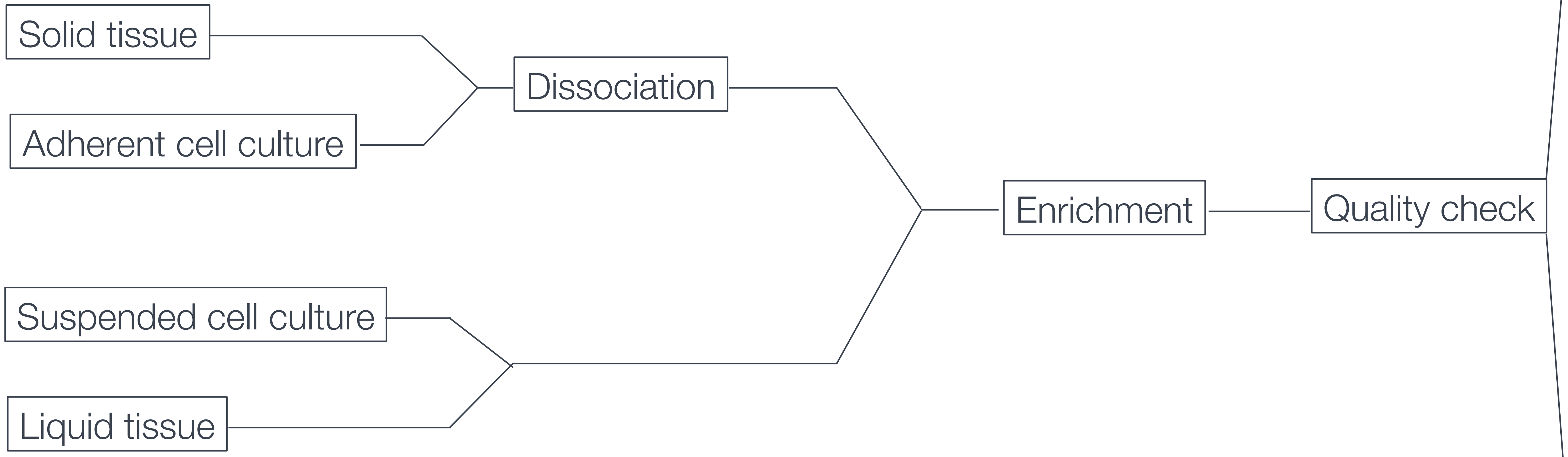
Reads/ cell barcode

Free-floating RNA

More background = More wasted reads

Key to Success: Sample Preparation

- 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?



Solid tissue

Adherent cell culture

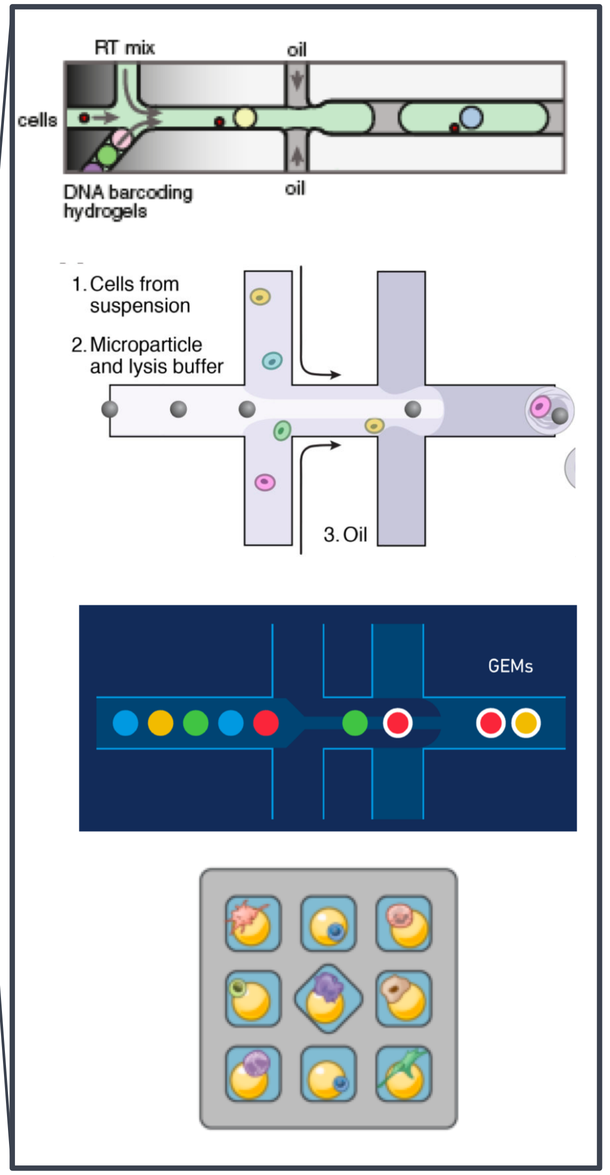
Dissociation

Suspended cell culture

Liquid tissue

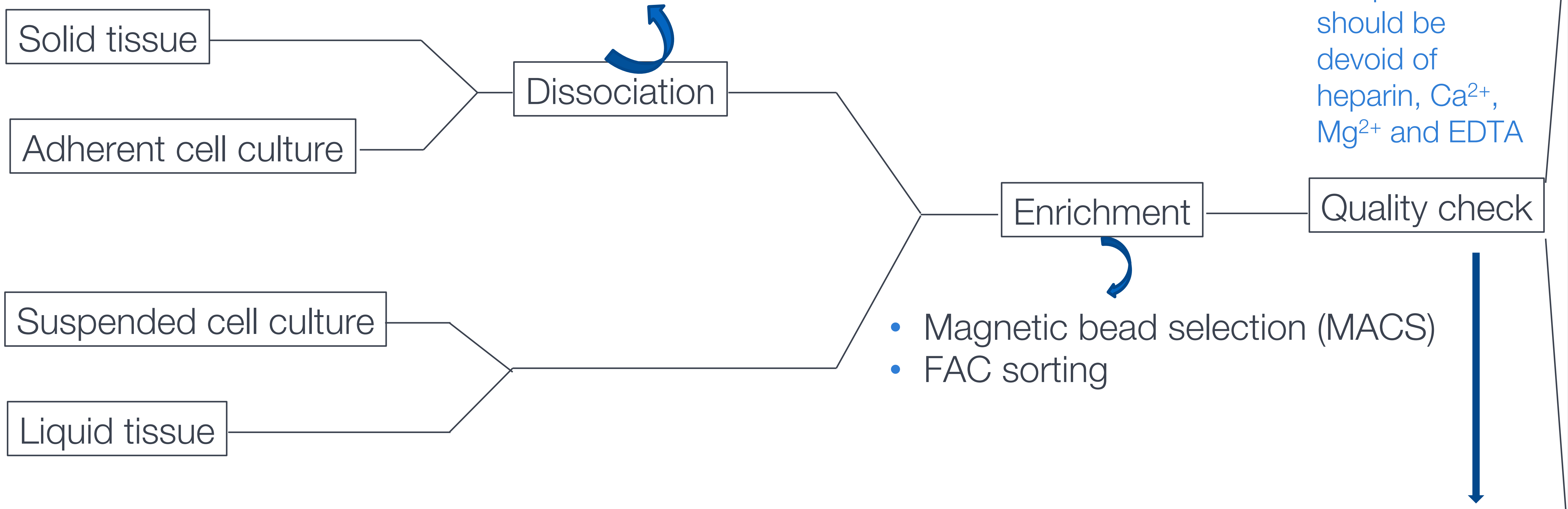
Enrichment

Quality check



Sample preparation protocol varies by cell-type

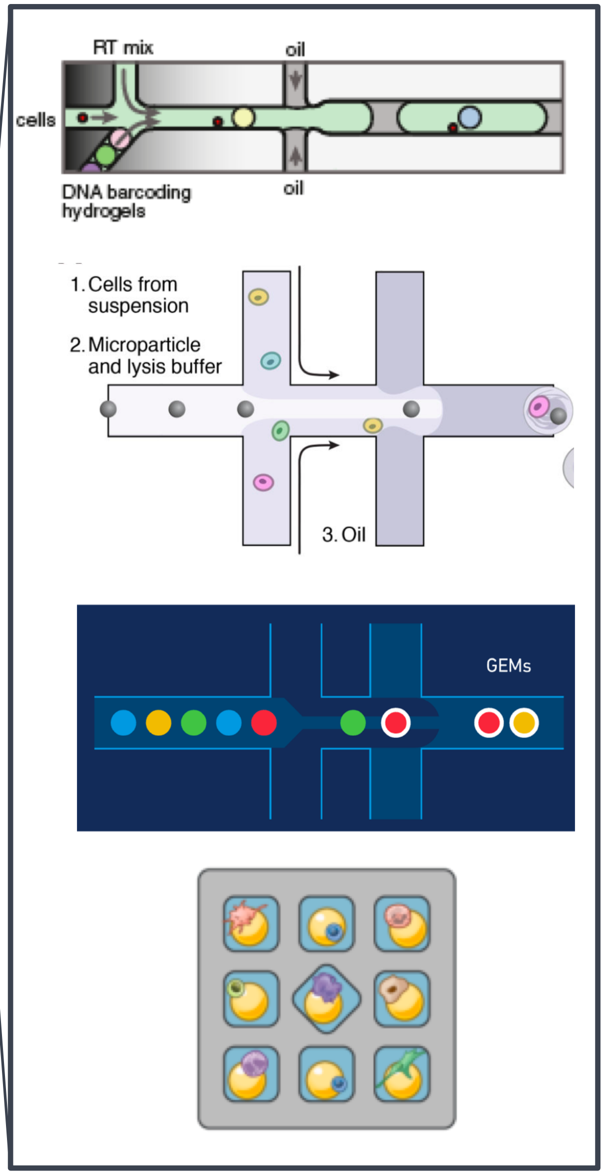
- Enzyme-based dissociation- trypsin, collagenase, liberase, accutase.
- Gentle washes.
- Dead cell removal kit, filtering out the debris.
- Density gradient (Ficoll, Optiprep)



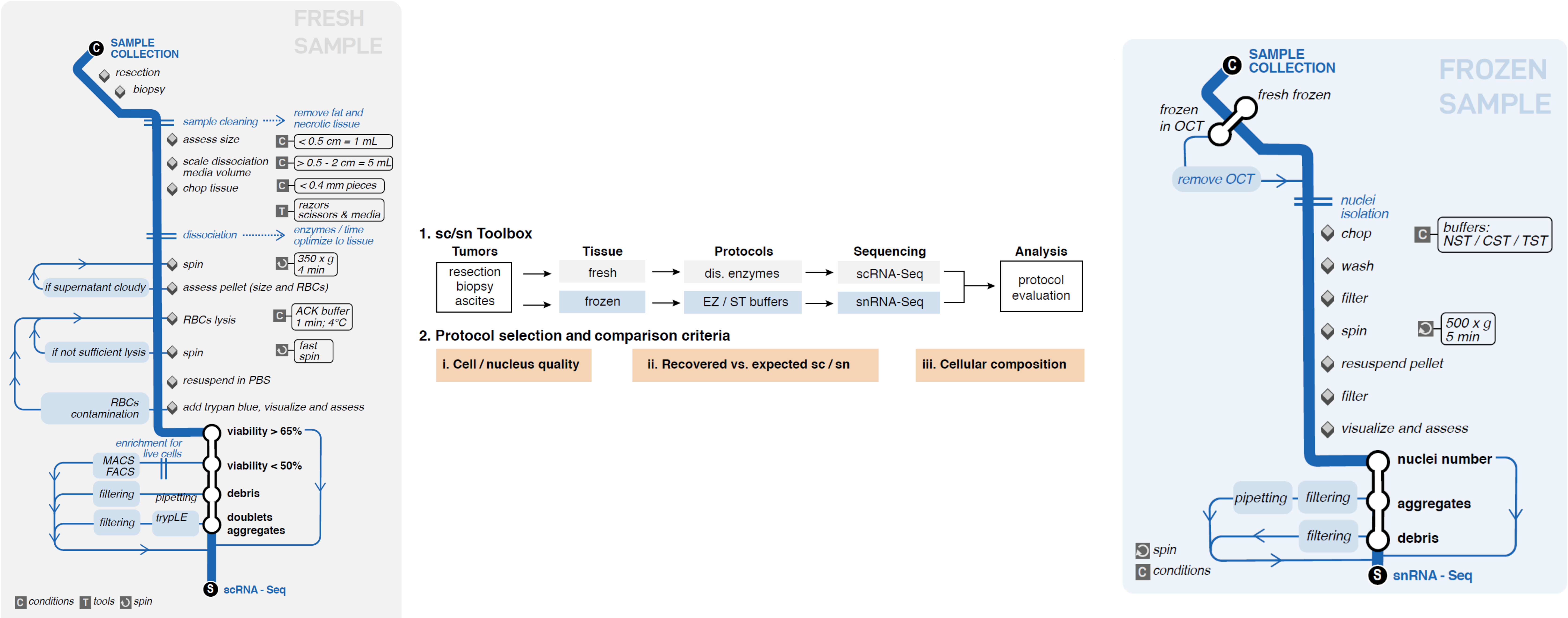
- Magnetic bead selection (MACS)
- FAC sorting

Note: the final sample buffer should be devoid of heparin, Ca^{2+}, Mg^{2+} and EDTA

- Cell concentration
- Cell Viability (Trypan Blue)



Sample Preparation Varies by Cell-Type



Slyper M., et. Al., A single-cell and single-nucleus RNA-seq toolbox for fresh and frozen human tumors. bioRxiv. doi: https://doi.org/10.1101/761429

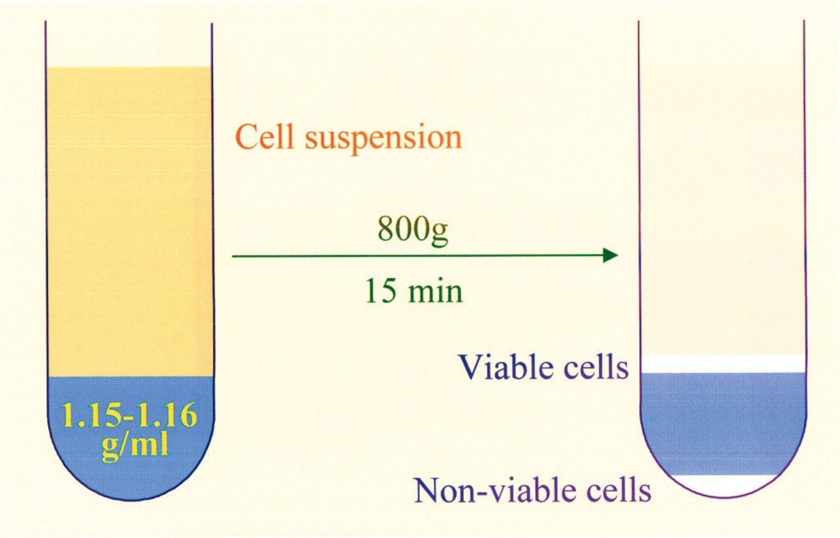
Sample Preparation: increasing cell viability

- Mild dissociation reagent (TrypLE, StemPro, Accutase, Liberase)
- Shorten dissociation time
- Reduce dissociation temperature
- Adding a DNAse step can help reduce clumping of dead cells.
- Using ROCK inhibitor/ apoptosis inhibitor (esp. epithelial cells)
- Avoid cell pelleting
- Avoid FACS sorting on more fragile cell types
- Try magnetic activated cell sorting (MACS)

MACS
Miltenyi Biotec

Sample Preparation: single cell suspension

- Use appropriate cell strainer to get rid off clumps.
- Use metabolically inert, non-toxic density gradient media to fractionate cells.
  - We routinely use 15% Opti-prep to keep cells in single cell suspension while loading sample, but the concentration might vary from cell to cell.



- Make sure final buffer does not contain calcium, EDTA, or heparin (inhibit RT).

Enrichment Methods: pros & cons

FACS (Fluorescence activated cell sorting)

Pros:

- Enrichment is robust.
- Yields good single cell suspension.
- Live/dead sorting by DNA stains, eg. DAPI.

Cons:

- Uses high pressure to sort the cells.
- Can introduce bias in the experiment.
- Long sample prep protocol.

MACS (Magnetic activated cell sorting)

Pros:

- Gentile on cells and faster protocol
- Greater number of cells can be processed.
- Not limited by FACS sorter availability.

Cons:

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

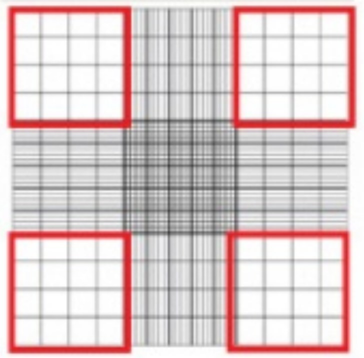
Enrichment Methods: cell staining

- Standard staining used for FACS
- DAPI staining to remove dead cells in sorting
- Live cell sorting: 7-AAD from Biolegend and Live 488 from Thermo
- Biotium’s NucSpot 470 for nuclear staining; requires cell permeation

Can always test if your dye of interest will inhibit RT in a mock tube based reaction.

Sample Preparation: cell numbers

- Droplet methods have a 10,000-25,000 cell minimum
  - need ~50-100 cells with a unique transcriptome to identify a population cluster
  - 100-1000 cells per ul = 100,000-1,000,000 cells per ml
- Count cells by hemocytometer – do not trust sorter counts
  - counts from the sorter are often ½ of actual cell counts
- Try negative selection to remove unwanted cells
- Sort on a broader marker to increase cell numbers
- For unavoidably low density samples
  - spike the sample with cells with distinct expression profile



Sample Preparation: buffers

- Suggest final sample prepared in:
    - PBS with 0.1-1% BSA
    - Defined media without calcium, magnesium, or EDTA
- 2% FBS in defined media without calcium or EDTA.
- Make sure buffer does not contain calcium, magnesium, EDTA, or heparin (inhibit RT).

Sample Preparation: viability checks

- Check viability of sample over time
  - If viability decreases over a short period of time this will be reflected in transcriptional data.
  - Will see high mitochondrial read counts.
- Check single cell suspension supernatant for the presence of free floating RNA (Ribogreen)
  - Creates background noise in all samples and complicates analysis.
- Number of trypan positive cells $\propto$ number of wasted sequencing reads
  - If 30% of your cells are dead at the time of encapsulation then at most you will be able to use 70% of your sequencing data.

Sample Preparation: dead cell removal

- FACS out dead cells
  - Will have all associated complications of FACS.
- Miltenyi dead cell removal kit
  - Magnetic beads used to remove dead cells & debris.

<u>Food For Thought</u>

- How many dead cells are you removing?
- What does this mean for the biology you are studying?
- Record your sample prep metadata!!!

Sample Preparation: cryopreservation

- Several papers have come out using various cryopreservation techniques on samples (PBMC's or cell lines).
- Success of cryopreservation is dependent on the sample type.
- Have seen this worked well on blood and immune cells.
- Key is the viability of the cells upon rehydration.
- Consider Nuc-seq as an option from cryopreserved cells.

Alles, J. et al. Cell fixation and preservation for droplet-based single-cell transcriptomics. BMC Biol. 2017 (doi: 10.1186/s12915-017-0383-5)
Guillaumet-Adkins, A. et al. Single-cell transcriptome conservation in cryopreserved cells and tissues. Genome Biol. 2017 (doi: 10.1186/s13059-017-1171-9)
Habib, N. *et al.* Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* 2017 (doi: 10.1038/nmeth.4407)

Sample Preparation: cryopreservation

- The quality of the tissue at the time of freezing is a major factor in the quality of data downstream.
- Single Cell Core has seen best results will cells frozen in their standard growth media supplemented with 5% DMSO.
  - Primary cells observed to have 20% cell death upon thawing.
- If freezing tissue for use later, you may want to consider cryopreserving 50 mg tissue chunks in BAM Banker cryopreservative. (Core has not tested yet)

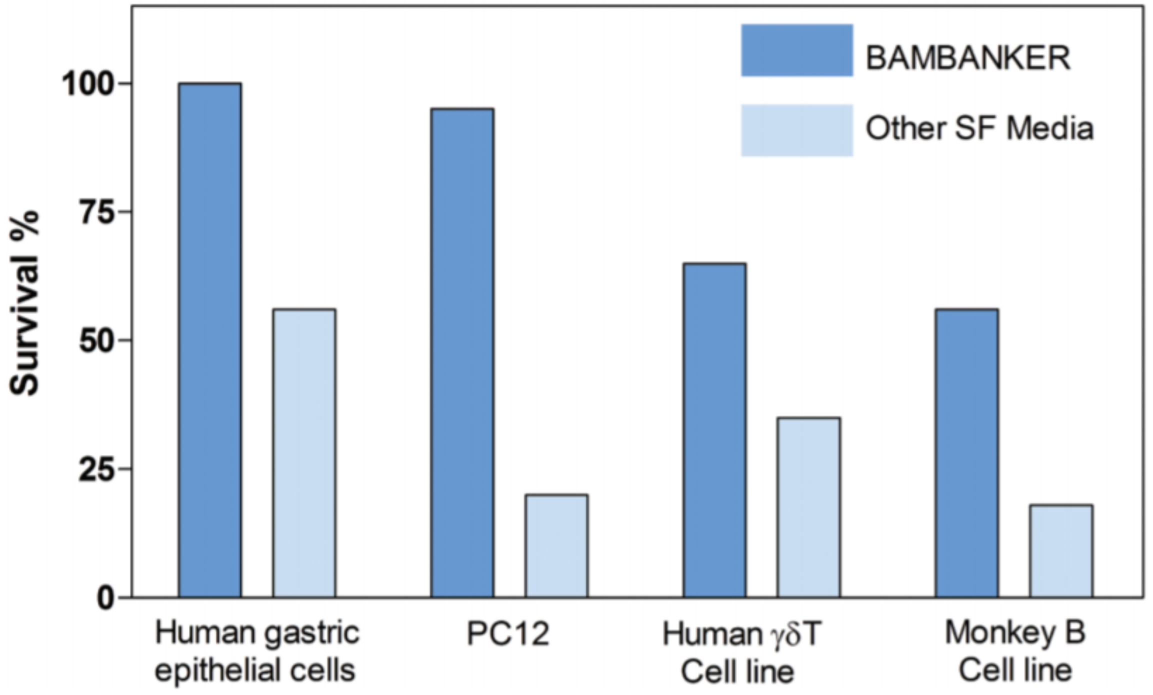
https://www.protocols.io/view/isolation-of-nuclei-from-frozen-tissue-for-atac-se-6t8herw/guidelines
Corces MR, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods. 2017 Oct;14(10):959–962.

Sample Preparation: cryopreservation

BAMBANKER cell freezing media has been tested with numerous cell types and shown to deliver excellent cell recovery rates.

Cryopreservation tests with human gastric epithelial cells, PC12 (rat-derived pheochromocytoma) cell line, activated human γδT Cells and Monkey B Cell lines show improved cell survival using BAMBANKER cell freezing media when compared to another serum-free media.

Cryopreservation with BAMBANKER™

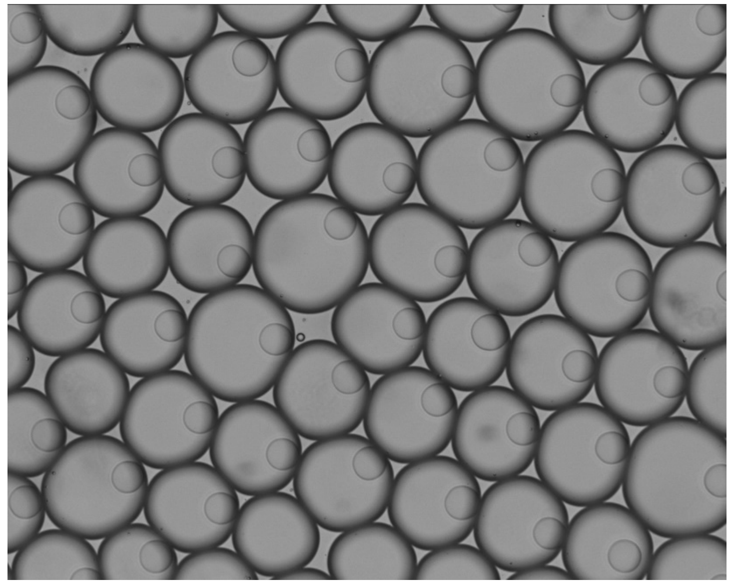


Also tested with Daudi (human B cell line), P3U1 (mouse myeloma cell line), K5625 (human leukaemia cell line), OKT4 (mouse hybridoma), activated lymphocytes derived from human peripheral blood and activated lymphocytes from mouse spleen.

BAMBANKER cell freezing media is available in two pack sizes to suit different usage patterns. For high volume users storing large batches of cells we recommend the 120ml pack size. Alternatively the pack of 5 smaller 20ml vials is ideal for individual use. To minimise cross contamination of cultures or culture facilities, separate vials could also be dedicated for use with specific cell lines or culture areas.

- Ready-to-use
- Freeze & store at -80 C
- Quality assured*

Sample Preparation: single nuclei RNA-seq

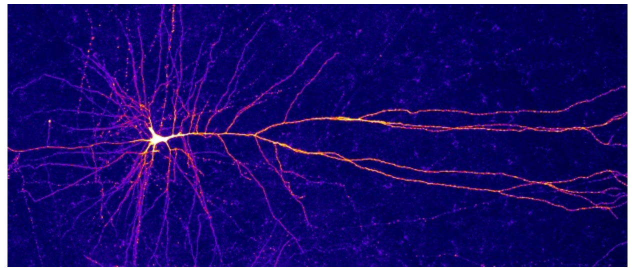


- Extract nuclei from sample of interest.
- Removes transcriptional noise from dead/dying cells.
- Most often used for neuronal samples.
- Good for flash frozen clinical samples.
- Several studies have shown nuclear transcripts represent a considerable portion of the whole cell transcripts.
- Analysis is more difficult due to presence of introns and non-coding RNA.

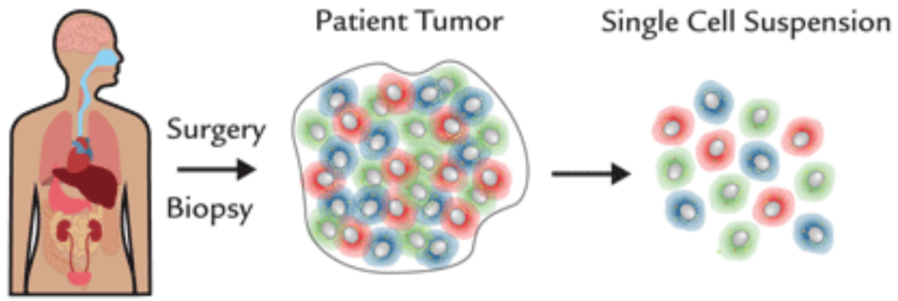
https://www.frontiersin.org/articles/10.3389/fcell.2018.00108/full#B12

Single nuclei RNA-seq

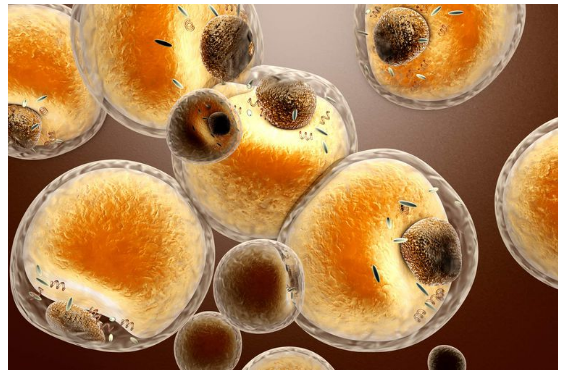
- Most often used for neuronal samples (difficult to isolate)



- Good for flash frozen clinical samples (low viability)

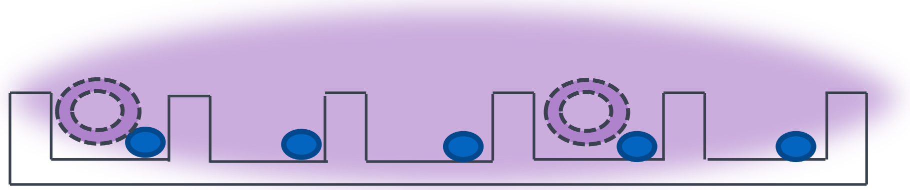


- Adipose tissue (fat inhibits RT in whole cell scRNA-seq)

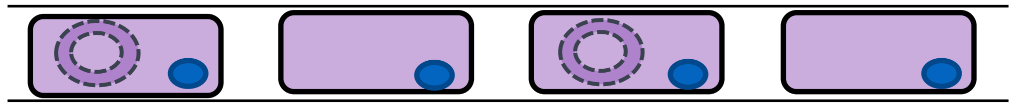


Sample Preparation: single nuclei RNA-seq

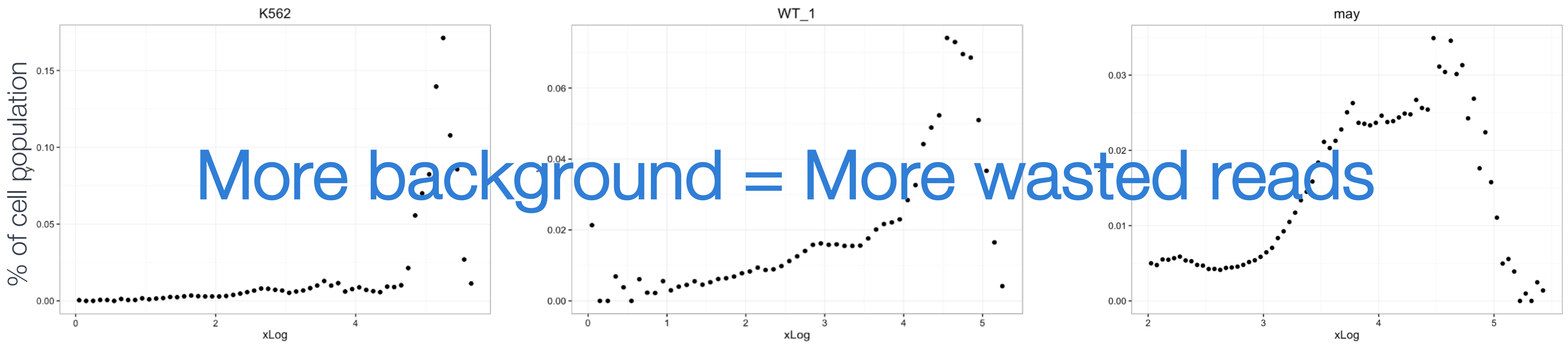
- Good single nuclei suspension. No clumps and minimal debris.
- Nuclear membrane integrity is required until cells are encapsulated.
- May need to DNase treat nuclear prep.



Nanowell-based barcoding platform



Droplet-based barcoding platform



More background = More wasted reads

Reads/ cell barcode

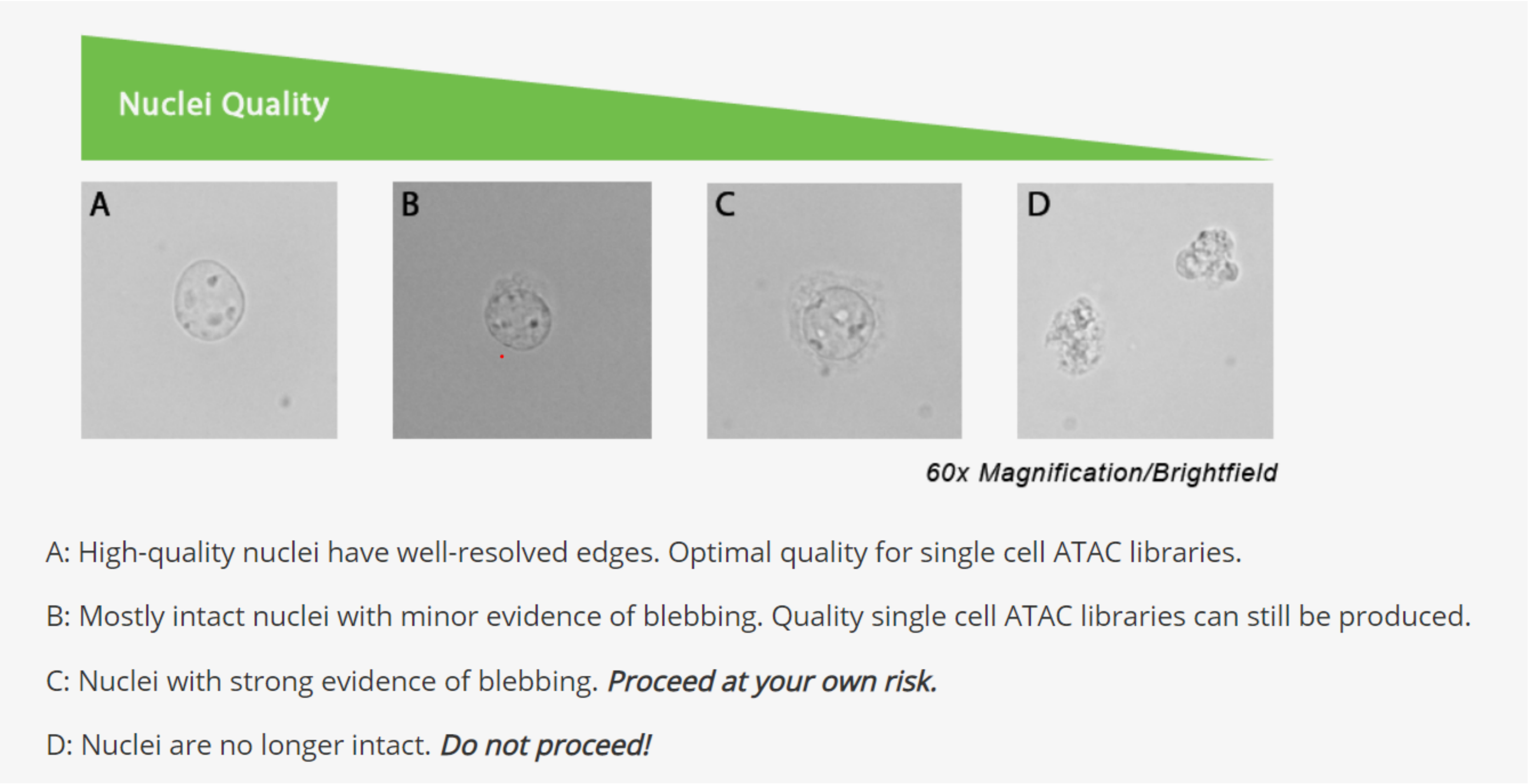
Reads/ cell barcode

Reads/ cell barcode

Ideal data

Free-floating ~~RNA~~ DNA

Sample Preparation: single nuclei RNA-seq



A: High-quality nuclei have well-resolved edges. Optimal quality for single cell ATAC libraries.

B: Mostly intact nuclei with minor evidence of blebbing. Quality single cell ATAC libraries can still be produced.

C: Nuclei with strong evidence of blebbing. ***Proceed at your own risk.***

D: Nuclei are no longer intact. ***Do not proceed!***

https://support.10xgenomics.com/single-cell-atac/sample-prep/doc/demonstrated-protocol-nuclei-isolation-for-single-cell-atac-sequencing

Single Nuclei Sample Preparation

‘Frankenstein’ protocol for nuclei isolation from fresh and frozen tissue for snRNAseq

Luciano Martelotto[1]

[1]University of Melbourne

Jun 01, 2019

2 *Works for me*

dx.doi.org/10.17504/protocols.io.3fkgjkw

Human Cell Atlas Method Development Community

Luciano Martelotto
University of Melbourne

BEFORE STARTING

All samples and reagents are kept on ice or at 4 °C (wet ice).

Prepare all buffers and reagents as described in the "Materials" section.

Tissue Homogenization

1 Mince/chop tissue with a razor blade to small pieces. The tissue may be as small as a grain of rice.

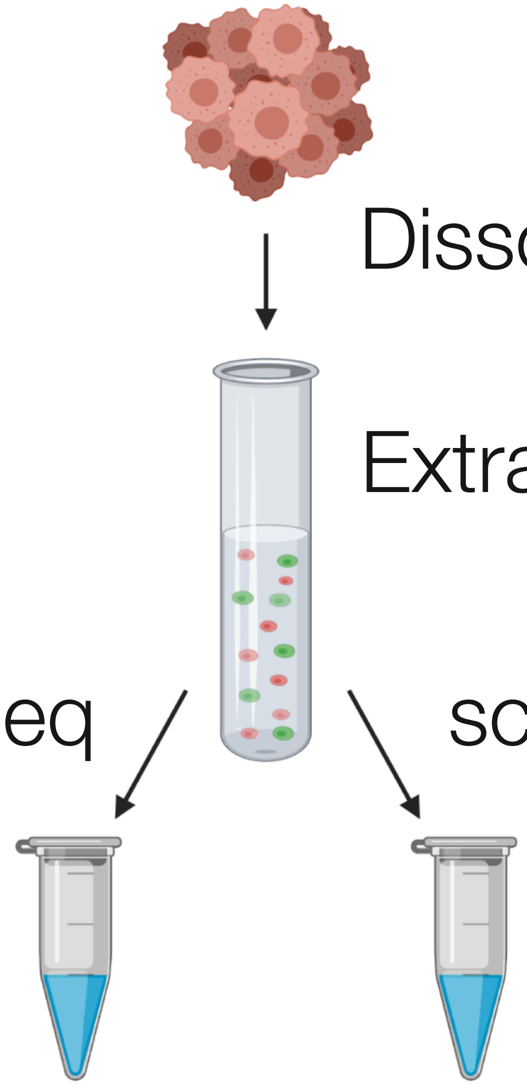
For mincing the tissue, you may take the tube out of ice, however, be quick and return to ice.

- Protocol has worked well for many users.
- Can use fluorescence activated nuclei sorting (FANS) to select for good nuclei.
- Suggested stains:

DAPI,7-AAD,DRAQ-7,SYTO

https://www.protocols.io/view/frankenstein-protocol-for-nuclei-isolation-from-f-3fkgjkw

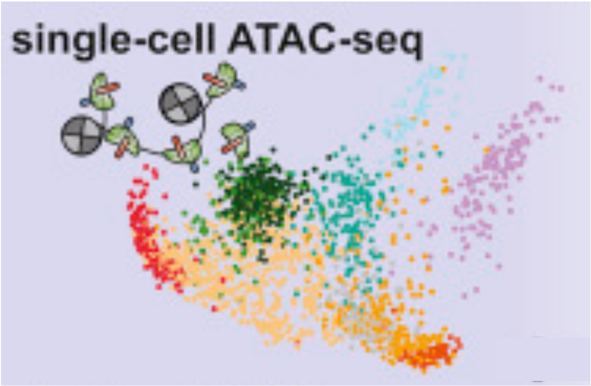
Nuclear Preparation



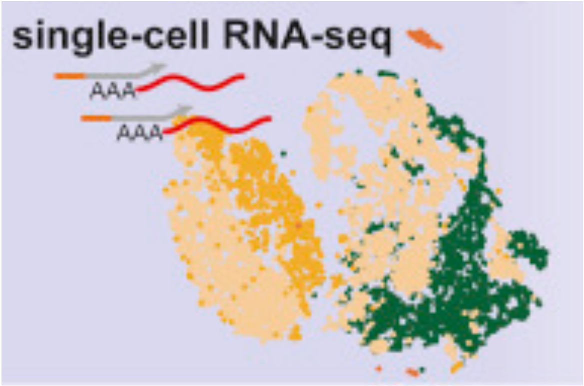
Dissociate Sample

Extract Nuclei

scATAC-seq

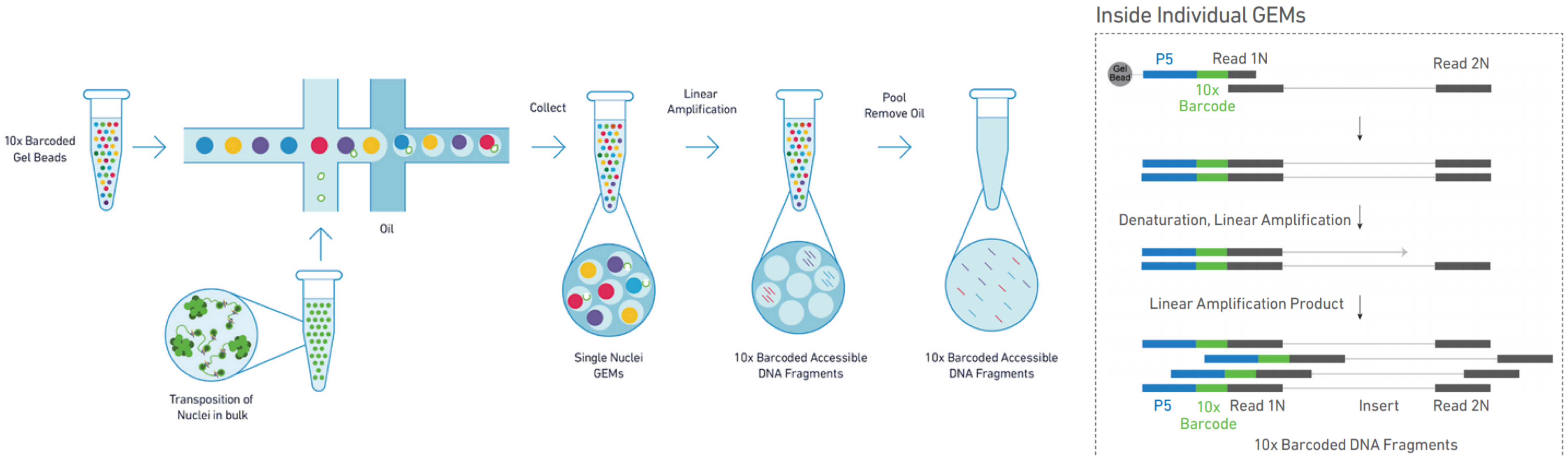


scRNA-seq



JD Buenrostro et al. Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation Cell. 2018 May 31;173(6):1535-1548.e16. doi: 10.1016/j.cell.2018.03.074.

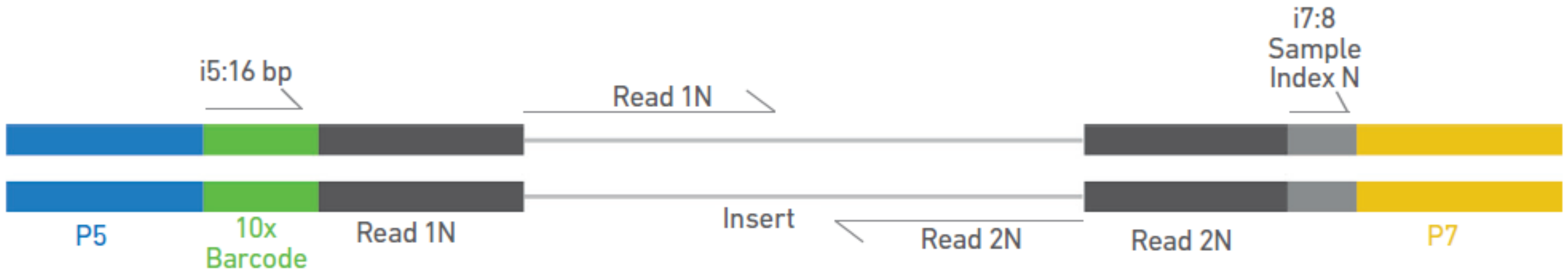
10x Genomics scATAC-seq Method Overview



~750,000 possible 10x barcodes

https://www.10xgenomics.com/solutions/single-cell-atac/

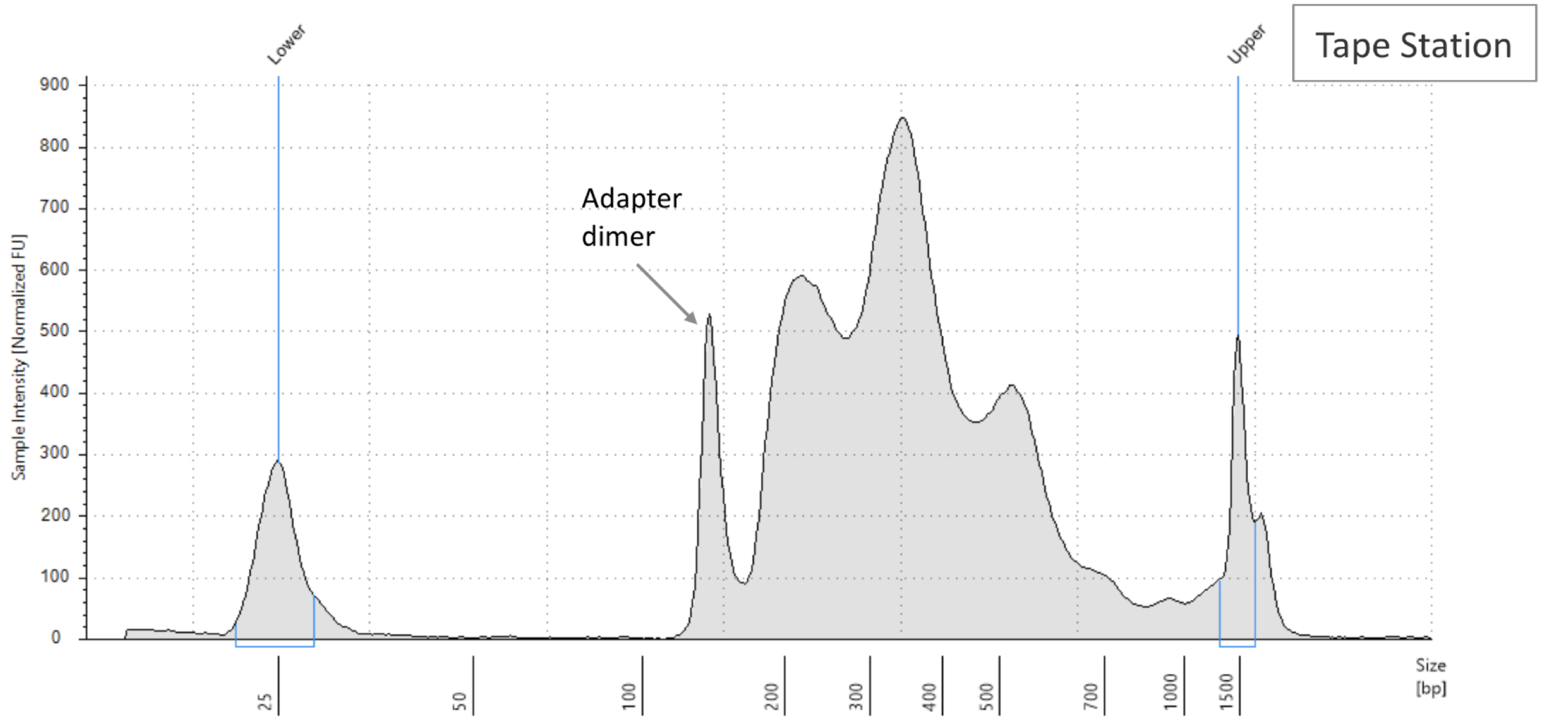
10x Genomics scATAC-seq Method Overview



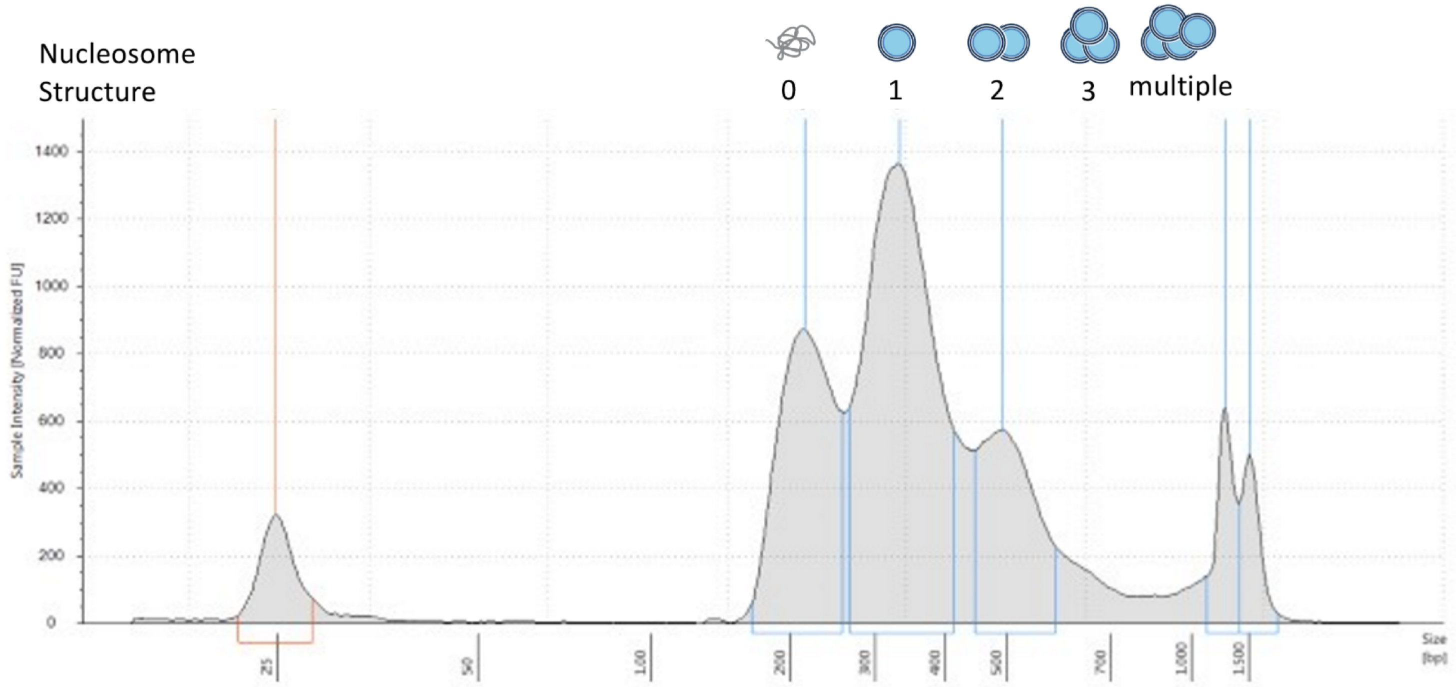
Sequencing Read	Description	Number of Cycles	Notes
Read1	Insert (genomic)	50	Can read longer if desired
I7 index	Sample Index	8	Sample Index
I5 index	Nuclei barcode	16	Nuclei barcode
Read 2	Insert (genomic)	50	Can read longer if desired

10x Genomics scATAC-seq Method Overview

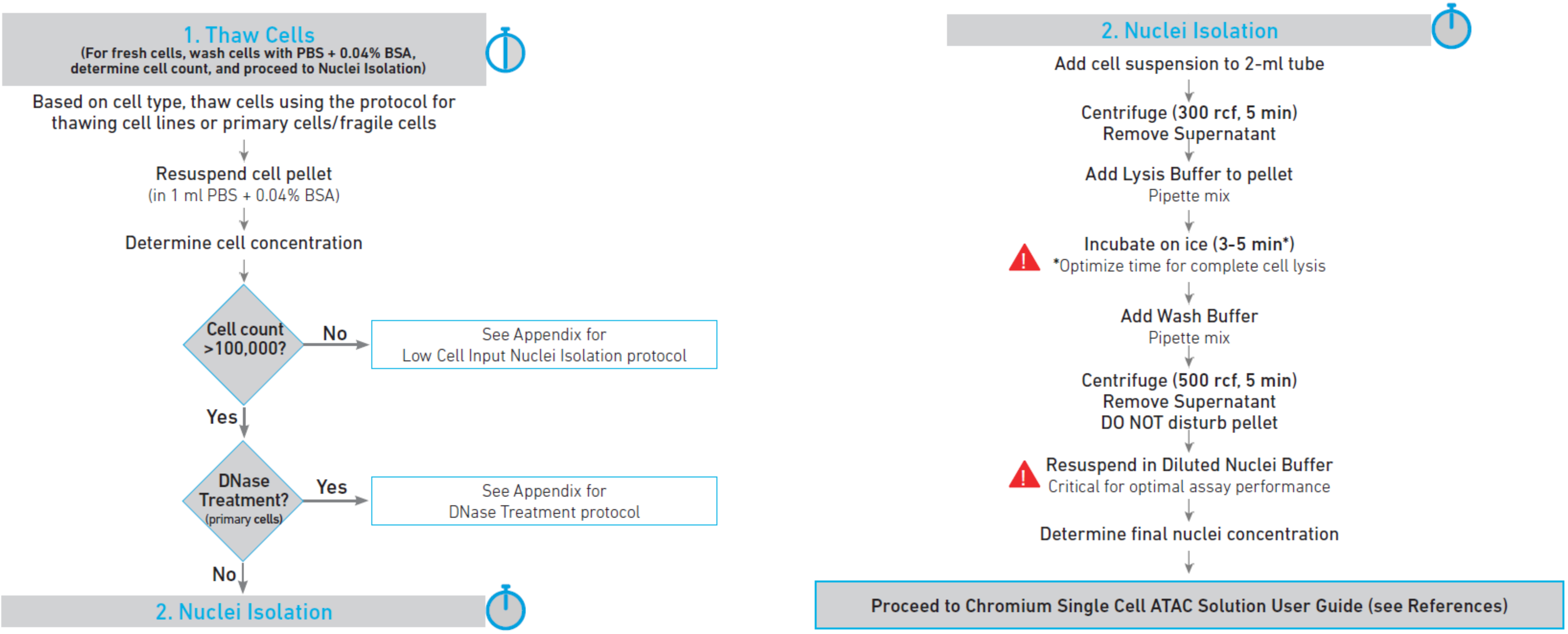
- 10x Example



- Single Cell Core



Sample Preparation: 10x crude nuclei protocol



1. Thaw Cells

(For fresh cells, wash cells with PBS + 0.04% BSA, determine cell count, and proceed to Nuclei Isolation)

Based on cell type, thaw cells using the protocol for thawing cell lines or primary cells/fragile cells

↓

Resuspend cell pellet
(in 1 ml PBS + 0.04% BSA)

↓

Determine cell concentration

↓

Cell count >100,000? — No → See Appendix for Low Cell Input Nuclei Isolation protocol

Yes ↓

DNase Treatment? (primary cells) — Yes → See Appendix for DNase Treatment protocol

No ↓

2. Nuclei Isolation

2. Nuclei Isolation

Add cell suspension to 2-ml tube

↓

Centrifuge (300 rcf, 5 min)
Remove Supernatant

↓

Add Lysis Buffer to pellet
Pipette mix

↓

⚠ Incubate on ice (3-5 min*)
*Optimize time for complete cell lysis

↓

Add Wash Buffer
Pipette mix

↓

Centrifuge (500 rcf, 5 min)
Remove Supernatant
DO NOT disturb pellet

↓

⚠ Resuspend in Diluted Nuclei Buffer
Critical for optimal assay performance

↓

Determine final nuclei concentration

↓

Proceed to Chromium Single Cell ATAC Solution User Guide (see References)

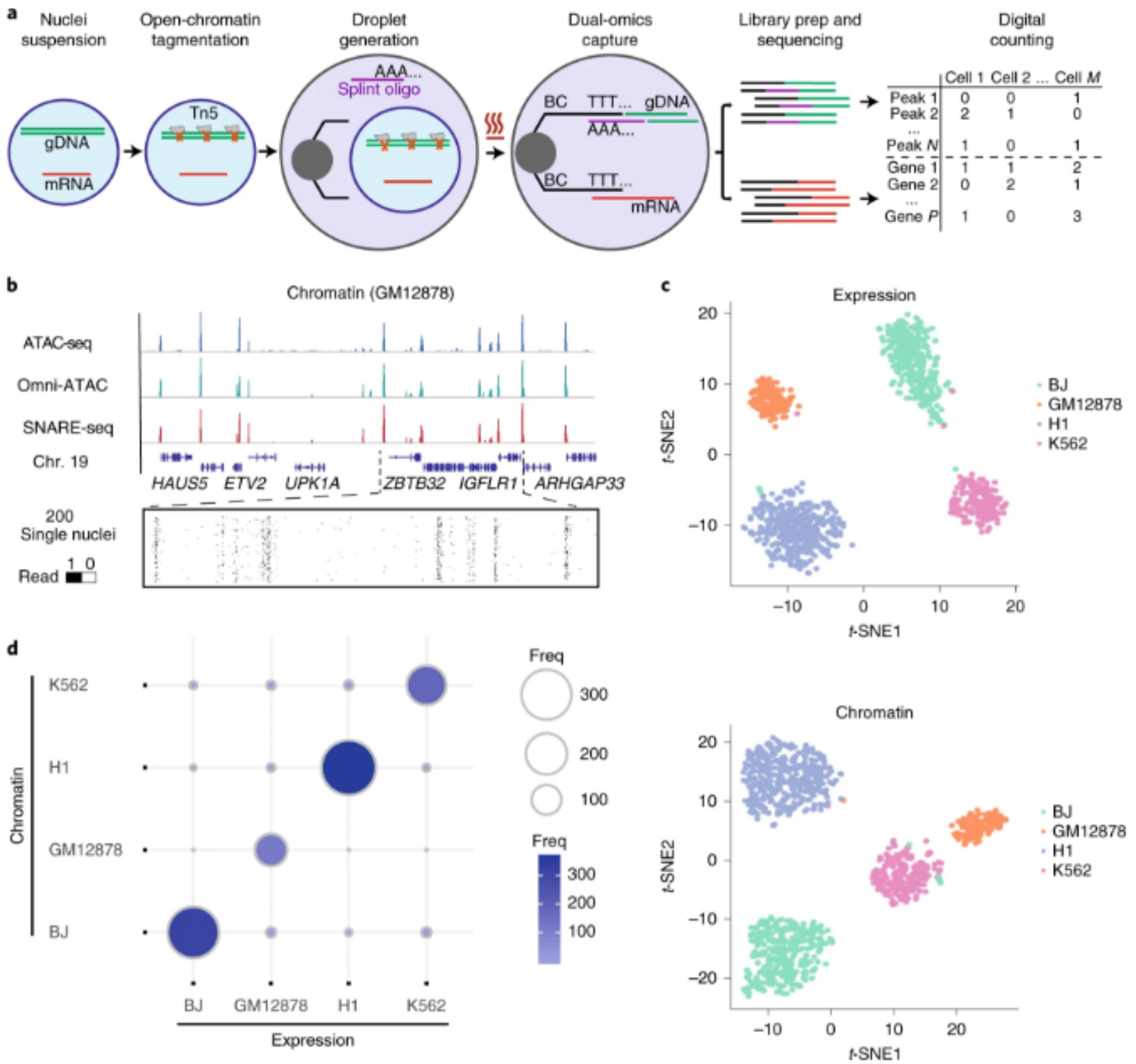
https://support.10xgenomics.com/single-cell-atac/sample-prep/doc/demonstrated-protocol-nuclei-isolation-for-single-cell-atac-sequencing

scATAC-seq Sample Preparation

- Test lysis time
- Assess level of free floating DNA
    - DNase treat to remove background DNA
    - Only for primary cells
- Visualize nuclei to asses clumping & debris
    - Use BSA up to 2% to minimize clumping
- Count nuclei – needs to be accurate

https://support.10xgenomics.com/single-cell-atac/sample-prep/doc/demonstrated-protocol-nuclei-isolation-for-single-cell-atac-sequencing

Single-nucleus chromatin accessibility and mRNA expression sequencing (SNARE-seq)

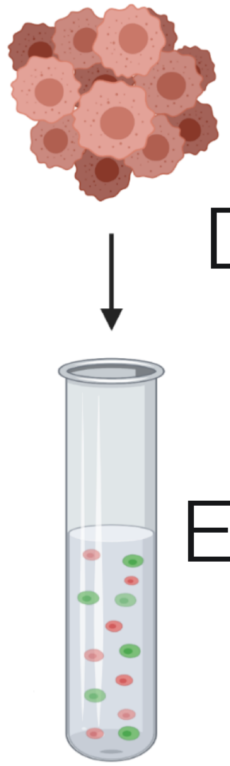


- Nuclei tagmentation /addition of polyA sequence.
- 10x Genomics releasing kit for this “in H2” (2^{nd} half of the year)
- Analysis will be challenging!

Chen et al. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell, Nature Biotech Dec 2019 (DOI: 10.1038/s41587-019-0290-0)

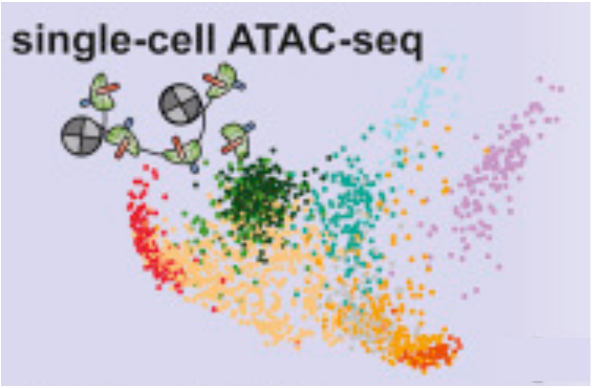
Dual scATAC-seq & snRNA-seq

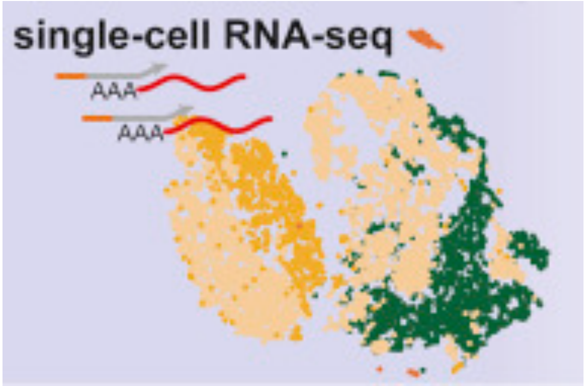
- Success of method highly dependent on sample quality!
- Need to strike the balance between over and under lysis.
- Will loose 50-60% of the nuclei in the sample processing.



Dissociate Sample

Extract Nuclei





Best practices to obtain high quality sample

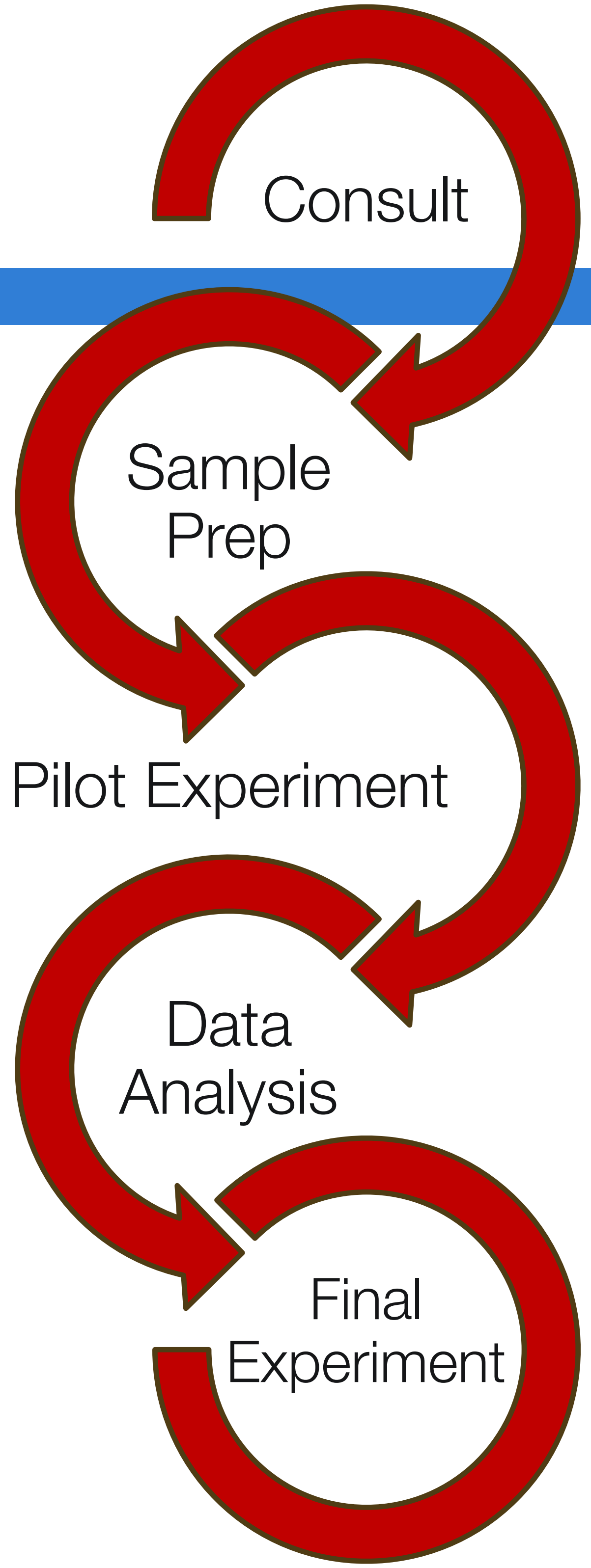
- Optimize a dissociation protocol that is best-suited for your cell type of interest.
- Short sample prep time.
- Maintaining low temperature.
- Gentle treatment
  - gentle lysis condition (low temp, short time)
  - short FACS time, slow sorting, bigger nozzle (in certain cases)
  - Gentle centrifugation (300-500xg) and resuspension
  - Removing debris by filter or density medium
- Include BSA (up to 1%) or FBS (up to 2%) in final buffer.

Resources for scRNA-seq Sample Prep

- https://www.protocols.io/
- https://support.10xgenomics.com/single-cell-gene-expression/sample-prep
- https://community.10xgenomics.com/

Single Cell Core Workflow

- Good sample prep is the key to success.
- A well planned pilot experiment is essential for evaluating sample preparation and for understanding the required number of cells.
- Do not rush to the final experiment.



Consult

Sample Prep

Pilot Experiment

Data Analysis

Final Experiment

Best Practices for Experimental Design

- Include biological replicates.
- Perform drug/treatment/model vs control on the same day.
- Randomize the order of samples run on different days.
- Use same sex littermates as controls in mouse experiments.
- Our experience is that library prep is the largest source of batch effect.
    - Collect all your samples in one study together then prep as one large library group.

Best Practices for Experimental Design

Pilot Study Experimental Design

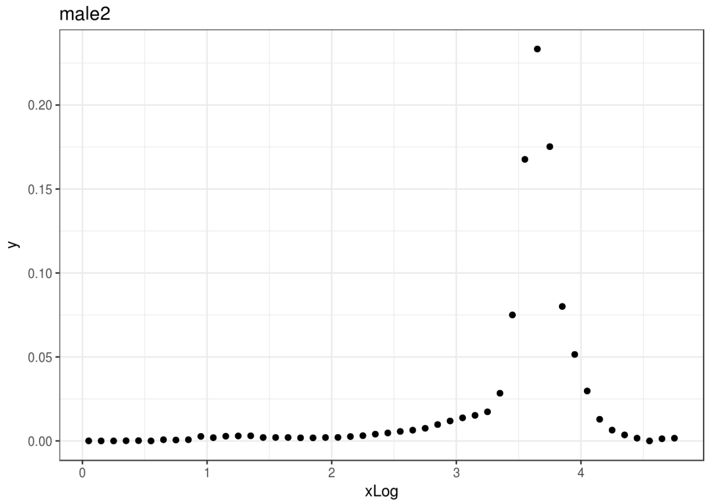
- Control vs diseased animal
- Each sample requires pooling several animals.
  - Control and diseased littermates pooled for a single sample.
- Both control and diseased samples run on same day.
- The entire experiment is repeated on a second day running samples.
  - Load samples in opposite order.
- Libraries from the four single cell samples are prepared as one batch.
- Sequencing and analysis performed
  - Check that sample prep was of good quality.
  - Determine that the desired information can be obtained from the experiment.

Introduction to Single Cell RNA Sequencing

- Common applications of single cell RNA sequencing.
- Overview of single cell RNA sequencing platforms.
- Modified scRNA-seq workflows
- Sample preparation and experimental design.
- Effects of sample prep and sample type on analysis.

How Sample Prep Effects Data

Ideal sample

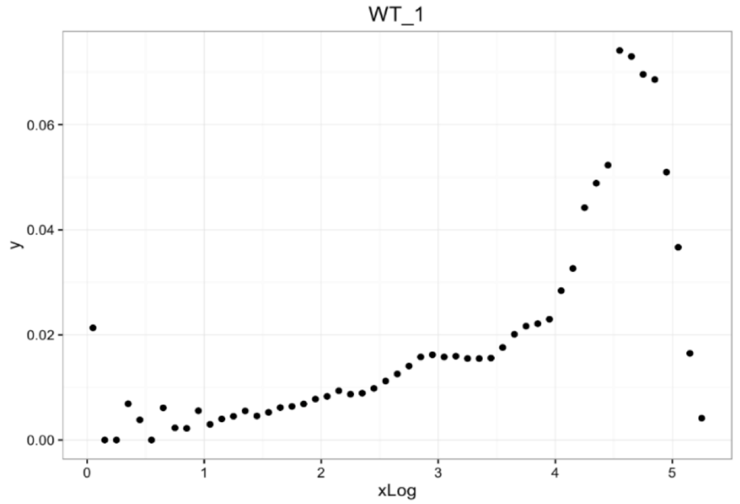


Proportion of population

Log10 reads per cell barcode

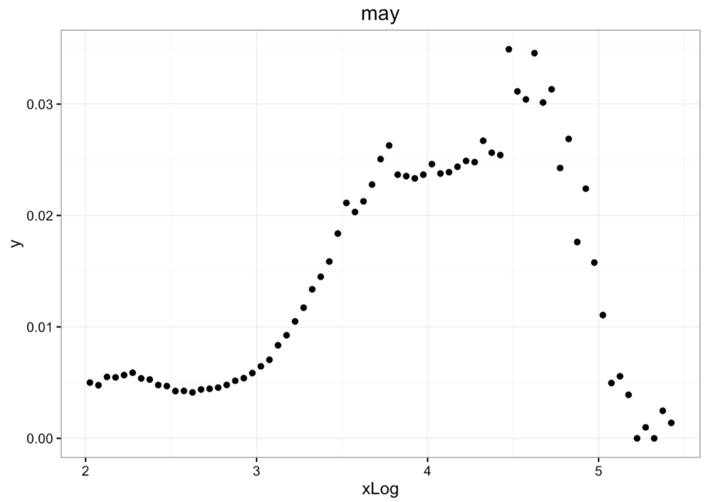
Messier Distribution

common



Failed Run

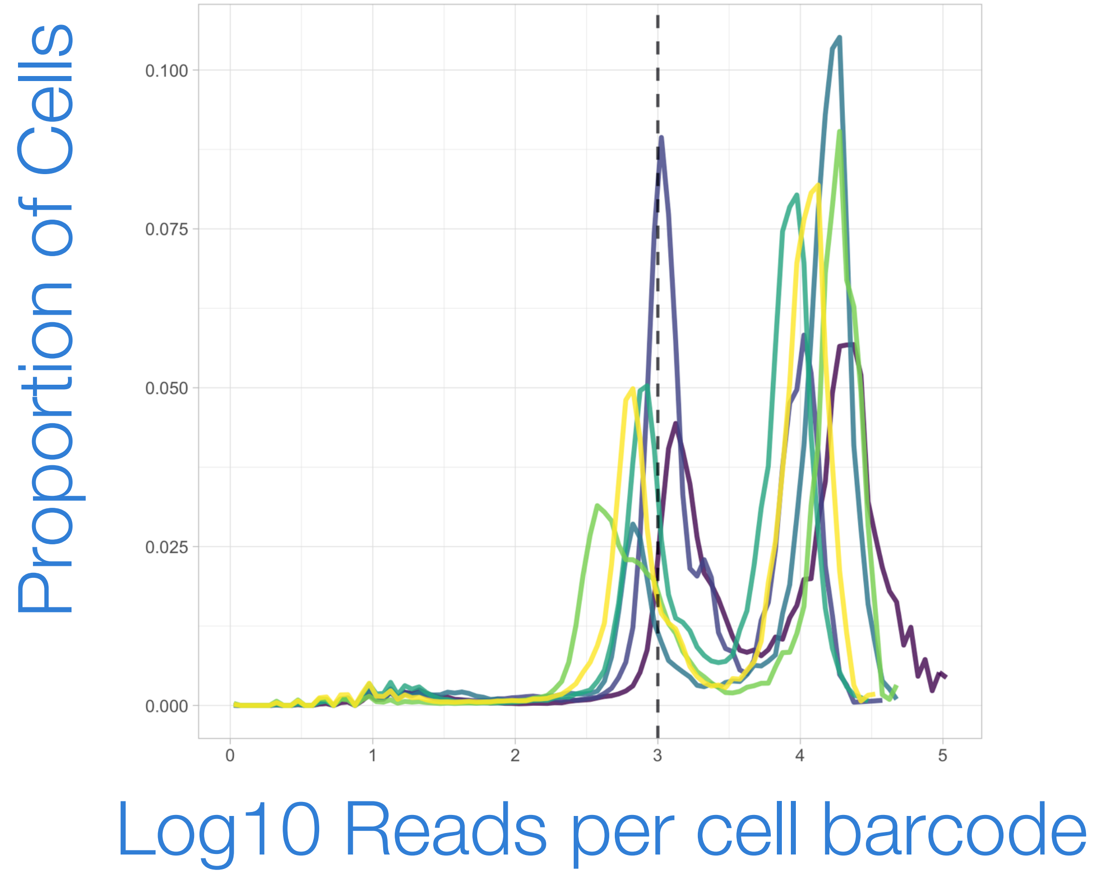
free floating RNA



- Cut off usually remove any cell with < 10,000 / 20,000 barcodes per cell
- It is normal for single cell RNA-seq data to contain a large number of low complexity barcodes.
- Exact threshold will depend on sample

Provided by the Harvard School of Public Health Chan Bioinformatics Core

How Sample Type can Effect Data



- Bimodal peak is due to sample type. (infiltrating immune cells in tumor)
- Lower peak can get filtered out in analysis.
- May want to analyzer each peak separately.

Provided by the Harvard School of Public Health Chan Bioinformatics Core

Data Analysis: Quality Control (QC) metrics

- Reads per cell: How many reads assigned to a given cell barcode
- UMI per cell: “Novelty” score looks for greater diversity genes per UMI
- Genes detected: Genes with a non-zero count measurement per cell
- Mitochondrial counts ratio: Biomarker for cellular stress

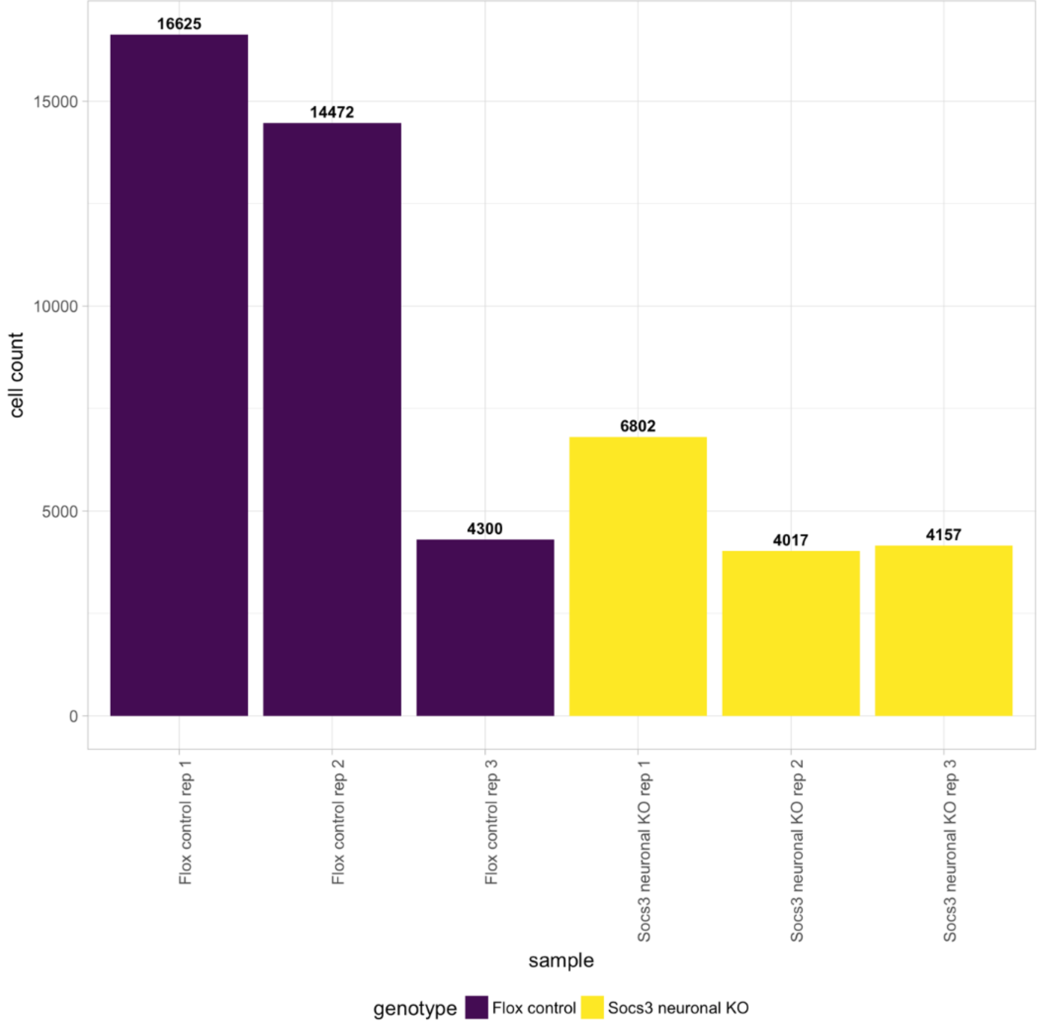
Filter Parameters (vary per experiment)

- > = 500 UMI counts per cell
- >= 500 genes per cell
- <=0.1% relative mitochondrial abundance
- >=0.8 novelty score

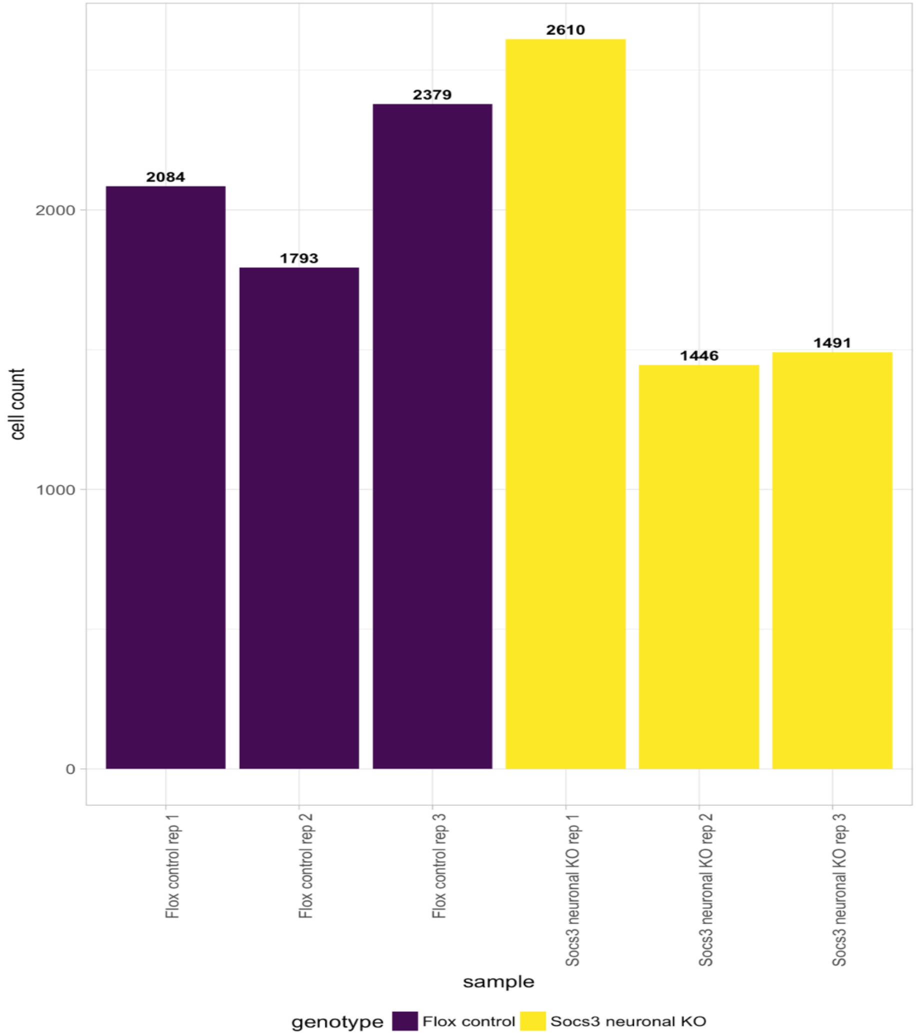
Provided by the Harvard School of Public Health Chan Bioinformatics Core

Data Analysis: filtering & correction

Pre-Filtering



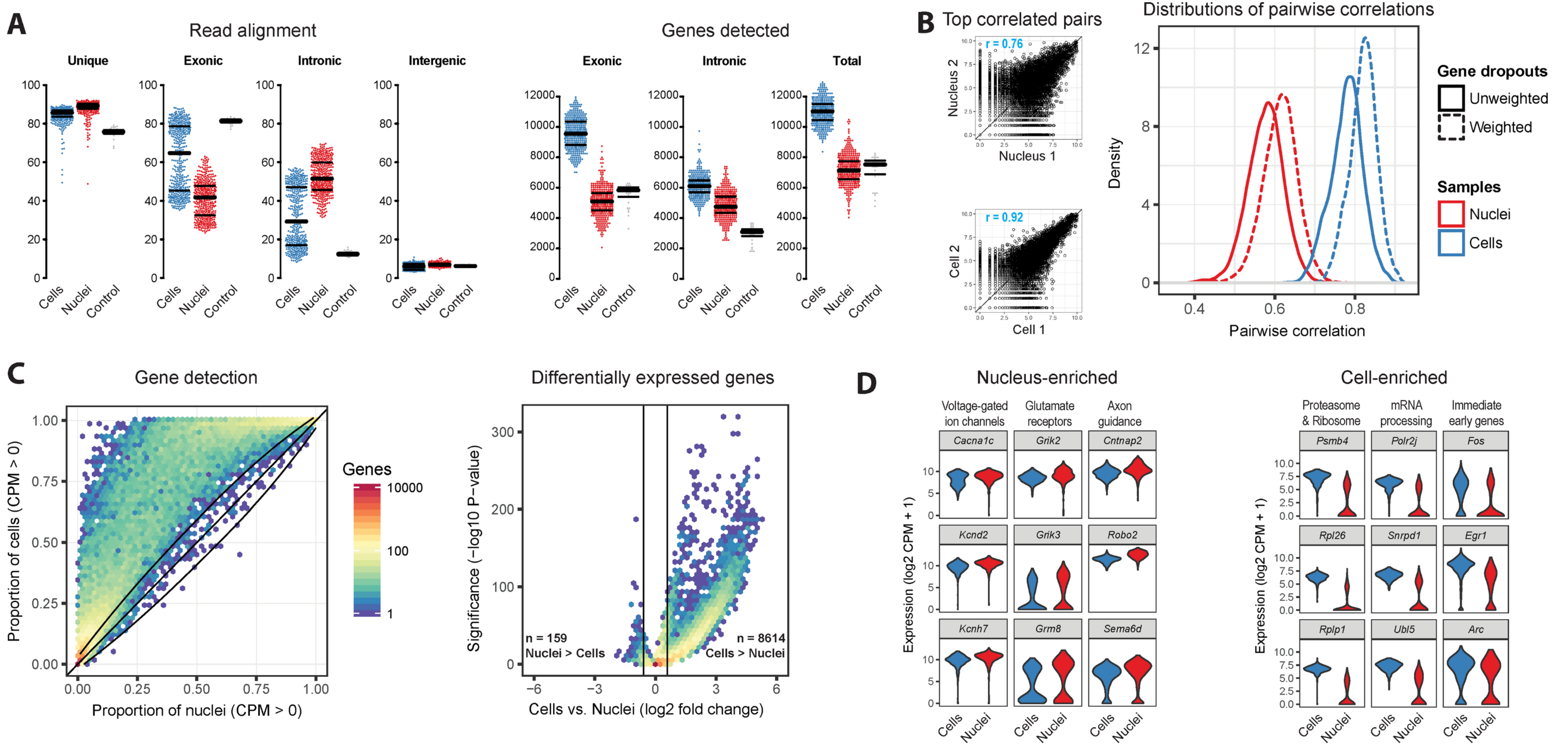
Post-Filtering



Libraries were of 3,000 cells. Post-filtering retains 50-80% of cells

Provided by the Harvard School of Public Health Chan Bioinformatics Core

Data Analysis: single cell vs. single nuclei



TE Bakken, et al., Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. PLoS One 13, e0209648 (2018)
https://doi.org/10.1371/journal.pone.0209648

Data Analysis: single cell vs. single nuclei

- Nuclei are more stable to expression changes from dissociation.
- Data from nuclei detects fewer genes per cell.
- Nuclei data has much higher percentage of intronic reads.
- Single nuclei sequencing seems to recapitulate the cell type classifications observed from single cell data.

TE Bakken, et al., Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. PLoS One 13, e0209648 (2018)
https://doi.org/10.1371/journal.pone.0209648

Final thoughts on scRNA-seq

- Practice your sample prep protocol. KEY to SUCCESS
- Start with a pilot sample set to ensure your protocol is working.
- Do not make your scRNA-seq run day the first day you run through the whole protocol.
- Be sure sequencing core understands the specific sequencing parameters needed for your scRNA-seq library.

qPCR

Precise quantitation is key to good clustering / sequencing