The Single Cell Sequencing Workflow:
A practical guideline and valuable insights to ensure experimental success

Arpita Kulkarni, Ph.D.

Associate Director, Single Cell Core, Harvard Medical School

arpita_kulkarni@hms.harvard.edu

Single Cell Core at HMS Quad

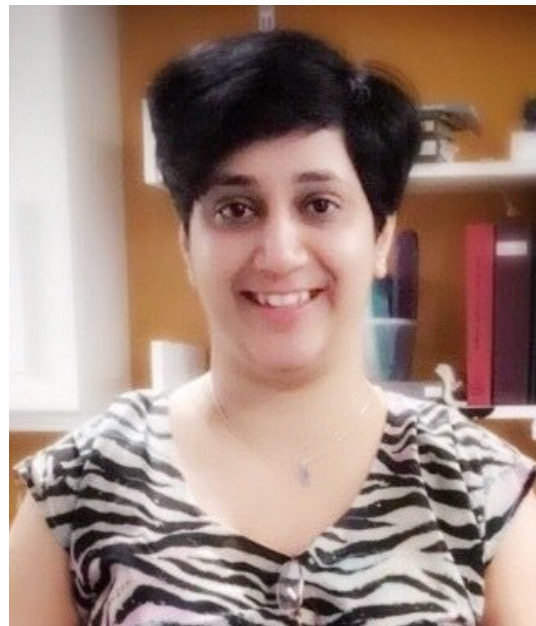We are here!

200 Longwood Ave., Armenise Room 517

https://singlecellcore.hms.harvard.edu

@HMS_SCC

Our Team

Mandovi Chatterjee, PhD
Director

Arpita Kulkarni, PhD
Associate Director

Theresa Torre
Research Technician

Allon Klein, PhD
Faculty Advisor

First name_Last name@hms.harvard.edu
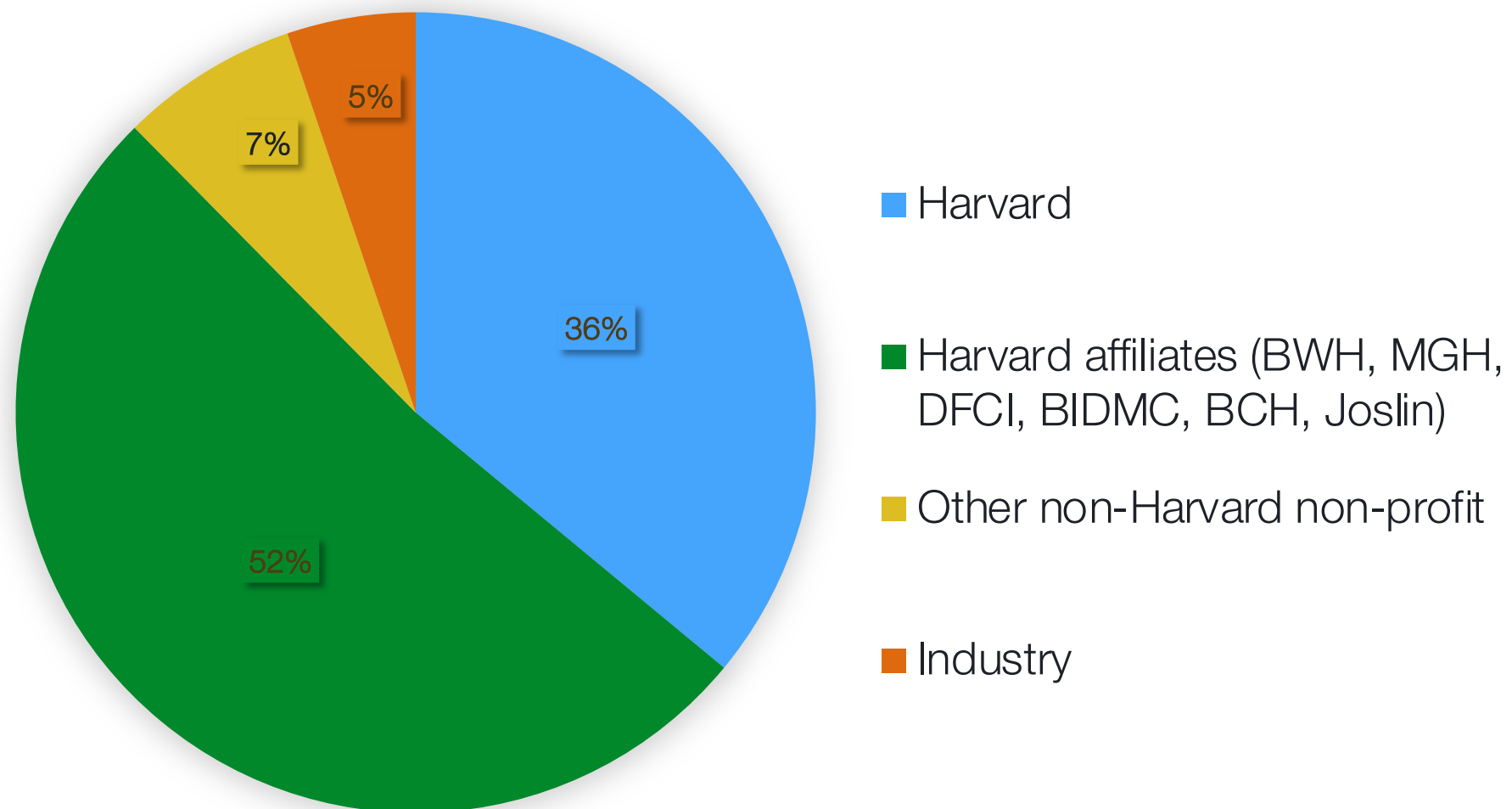
Single Cell Core at HMS Quad

Mission: Enable novel discoveries by assisting in the design, execution, and interpretation of single cell genomics assays using state-of-the-art tools

The Single Cell Core

We house different high-throughput platforms that allow encapsulation, barcoding, & library prep from single cells for single cell/ spatial transcriptomics and epigenomics

Outline for today's talk

- scRNAseq vs bulk (which one to use when?)

- scRNAseq workflow -

Step I	Step II	Step III	Step IV

Tissue preparation	Single cell isolation and library prep	Sequencing and primary analysis	Data visualization and interpretation
Mechanical, enzymatic, or combinatorial methods with enrichment (optional) and QC	Low-throughput: microdissection, FACS High-throughput: microfluidics, droplet-based, microwell approaches	Illumina Sequencing System	Multiple commercial and freeware secondary and tertiary analysis packages available

We know tissues are heterogeneous

Adipose
Fat
Connective tissue

Band Neutrophil
Normal Peripheral Blood
Eosinophil
Segmented Neutrophil

Small Intestine Mucosa
Epithelial cells
Goblet cells
Lamina propria
Muscularis mucosa (smooth muscle)

10x GENOMICS

Bulk RNA Sequencing (est. ~2000s)

Bulk RNA-seq
(masks cellular heterogeneity)

average expression level across a population

- comparative transcriptomics
- disease biomarker
- homogenous systems
- Great for studying broad level differences

Bulk RNAseq

→ mRNA-only libraries (polyA)

→ Whole transcriptome libraries
(all RNA species except rRNA)

Bulk RNA Sequencing (est. ~2000s)

Bulk RNA-seq
(masks cellular heterogeneity)

average expression level across a population

- comparative transcriptomics
- disease biomarker
- homogenous systems
- Great for studying broad level differences

Sometimes averages are not useful!

"Say you are standing with one foot in the oven, and the other foot in an ice bucket. According to the percentage people, you should be perfectly comfortable"
- Bobby Bragan

Single Cell RNA Sequencing

scRNA-seq
(uncovers cellular heterogeneity)

Population 1
Population 2
Population 3
Population 4

- define heterogeneity
- identify rare cell population(s)
- cell population dynamics

Captures cell to cell variation in gene expression

scRNAseq → mRNA-only libraries (polyA)
Single cell platforms not good at capturing RNA species other than mRNA!

The main difference between bulk and scRNA-seq is that in the latter each sequencing library represents a single cell, instead of a population of cells

Single cell vs Bulk RNA Sequencing: not an either/or situation

Which technique to use when?

Does your biological Q need scRNAseq?

RNA-Seq

Average expression level
- Comparative transcriptomics
- Disease biomarker
- Homogenous systems

scRNA-Seq

Separate populations
- Define heterogeneity
- Identify rare cell populations
- Cell population dynamics

CANCER RESEARCH UK | CAMBRIDGE INSTITUTE

Data Quality -Transcriptome Coverage (mRNA)

"Bulk RNAseq"

mRNA: TruSeq RNA-Seq (gold standard)

- Higher starting RNA material (500ng-)

- >~20,000 transcripts per cell

  - More when consider splice variants / isoforms

- Capture >80-95% of transcriptome depending

on sequencing depth

Data Quality -Transcriptome Coverage (mRNA)

"Bulk RNAseq"

mRNA: TruSeq RNA-Seq (gold standard)

- Higher starting RNA material (500ng-)

- >~20,000 transcripts per cell

  - More when consider splice variants / isoforms

- Capture >80-95% of transcriptome depending

on sequencing depth

"Single Cell Methods"

- Lower starting RNA (noisier gene expression) 10^3-10^6 cells

- 200 -10,000 transcripts per cell

- Capture <10-40% of the transcriptome

- less reads/cell → many transcripts (80-90%) will show up with zero counts in every cell (eg. GAPDH, ACTB)

scRNAseq can be very powerful but you want to be sure that it is the best method for your Q

Single cell vs Bulk RNA Sequencing: The face-off

SINGLE-CELL SEQUENCING VS. BULK SEQUENCING

PRICE
More expensive — Costs less

INTENSITY
Labour intensive — Less labourious

TYPE OF OUTPUT
More detailed information of individual cells — Measures average expression across a population of cells

EXPRESSION
Reveals cellular heterogeneity & subpopulation expression — Cellular heterogeneity is masked

STATISTICAL COMPLEXITY
More complex — Faster & less complex analysis

Common Applications of scRNA-seq

a) "cell atlas"-type studies
- Heterogeneous populations

b) "timeseries"-type studies
- Snapshots in biol. process

c) "screening"-type studies
- Single cells as individual expt.

Uncover cellular
heterogeneity

e.g. mouse brain atlas,
Tumor environment etc

Bio. trajectories,
Dev timelines,
lineage tracing

e.g. embryogenesis

Uncover GEX diff on
purturbation

e.g. CRISPR studies

"-omics" one can study at single cell level

Single cell _____

- Transcriptomics

- Epigenomics

- Genomics

- Proteomics

- Metabolomics

- Microbiomics

- Lipidomics

- Glycomics

- Multiomics

Each "–omics" produces large data

BUT
Integrating big data from multi "–omics"
presents a considerable statistical challenge

Spatial transcriptomics at (sub)cellular res

High resolution spatial profiling of scRNA expression in their native context

- All methods are in their infancy –

Sequencing or Imaging based

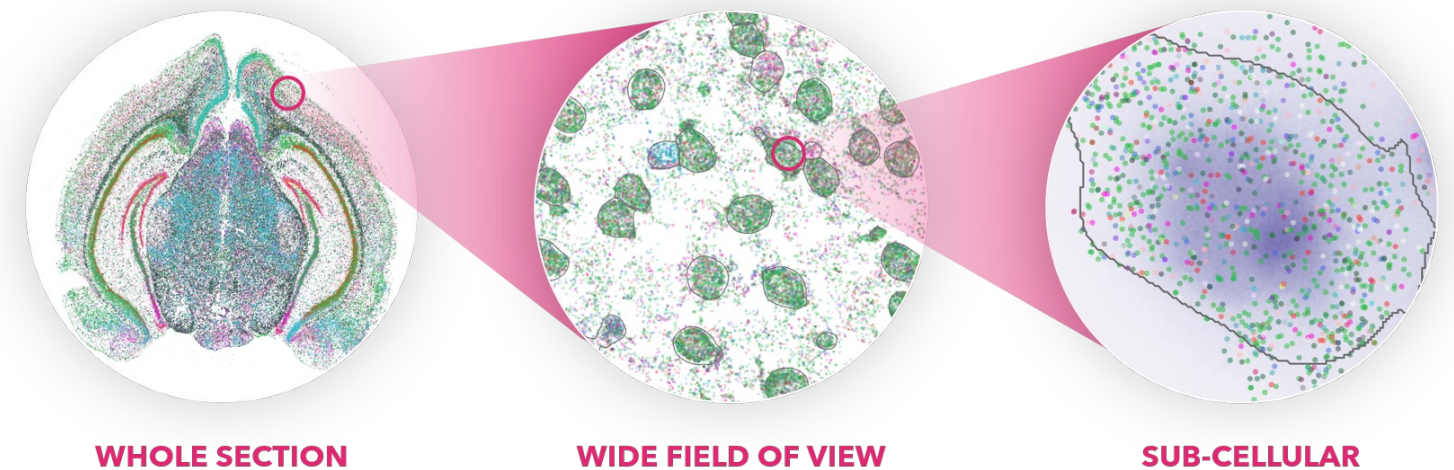- Require fresh-frozen tissue sections

Examples:

10x's Visium (50-100μm resolution), Xenium
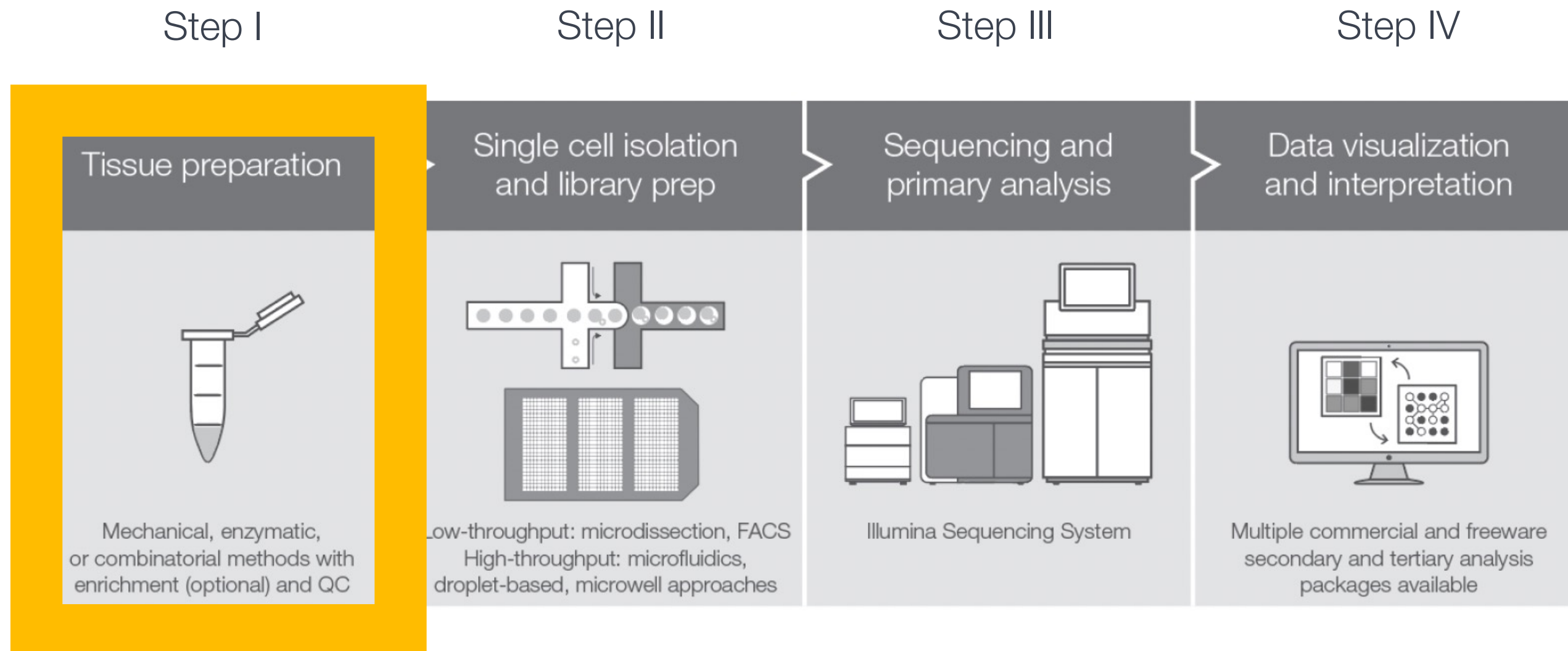
Vizgen's MERSCOPE

Nanostring's GeoMx, CosMx

vizgen

MERFISH • Multiplexed Error Robust Fluorescence In-Situ Hybridization
Profiling 483 genes with subcellular resolution across a full mouse coronal slice

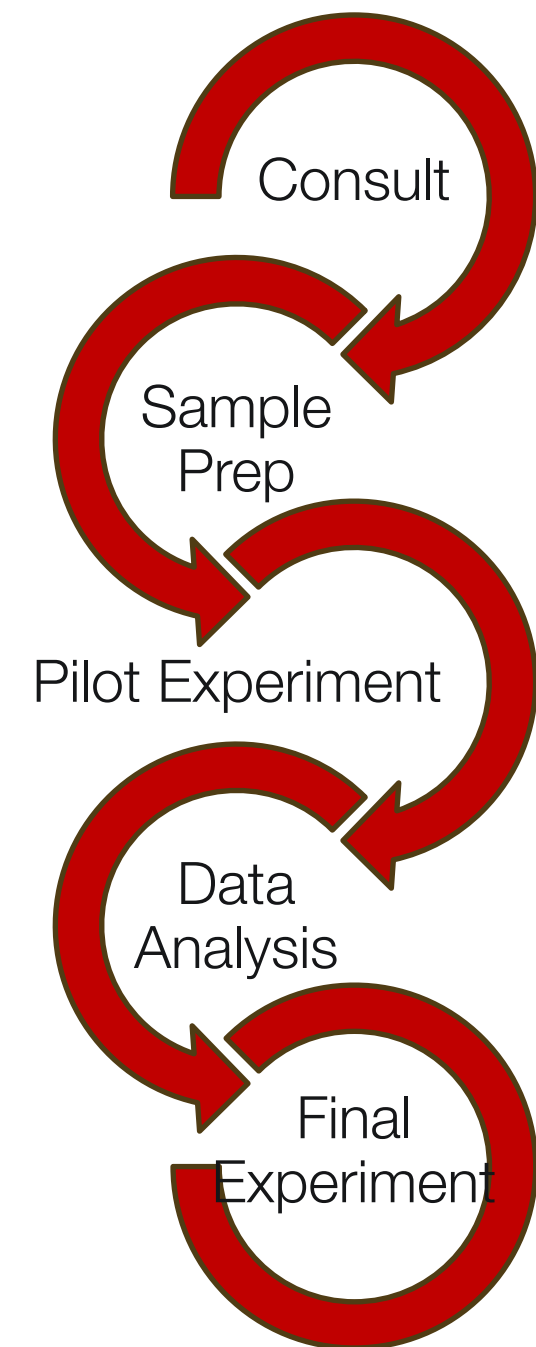WHOLE SECTION   WIDE FIELD OF VIEW   SUB-CELLULAR

SCC offers spatial services

Single Cell Sequencing Workflow – STEP I

Step I   Step II   Step III   Step IV

Step I — Tissue preparation: Mechanical, enzymatic, or combinatorial methods with enrichment (optional) and QC

Step II — Single cell isolation and library prep: Low-throughput: microdissection, FACS. High-throughput: microfluidics, droplet-based, microwell approaches

Step III — Sequencing and primary analysis: Illumina Sequencing System

Step IV — Data visualization and interpretation: Multiple commercial and freeware secondary and tertiary analysis packages available

Goal: Get high quality, viable, single cell suspension from tissue

STEP I – Pilot experiment planning

- Do not rush to the final experiment

- A well-planned pilot experiment is essential for
  - ✓ coming up w/ well defined bio. objectives
  - ✓ rational expt design/optimal approach for research Q
  - ✓ evaluating sample preparation
  - ✓ figuring our the required number of cells needed to answer your biological question

- Good sample prep is the key to success

Consult

Sample Prep

Pilot Experiment
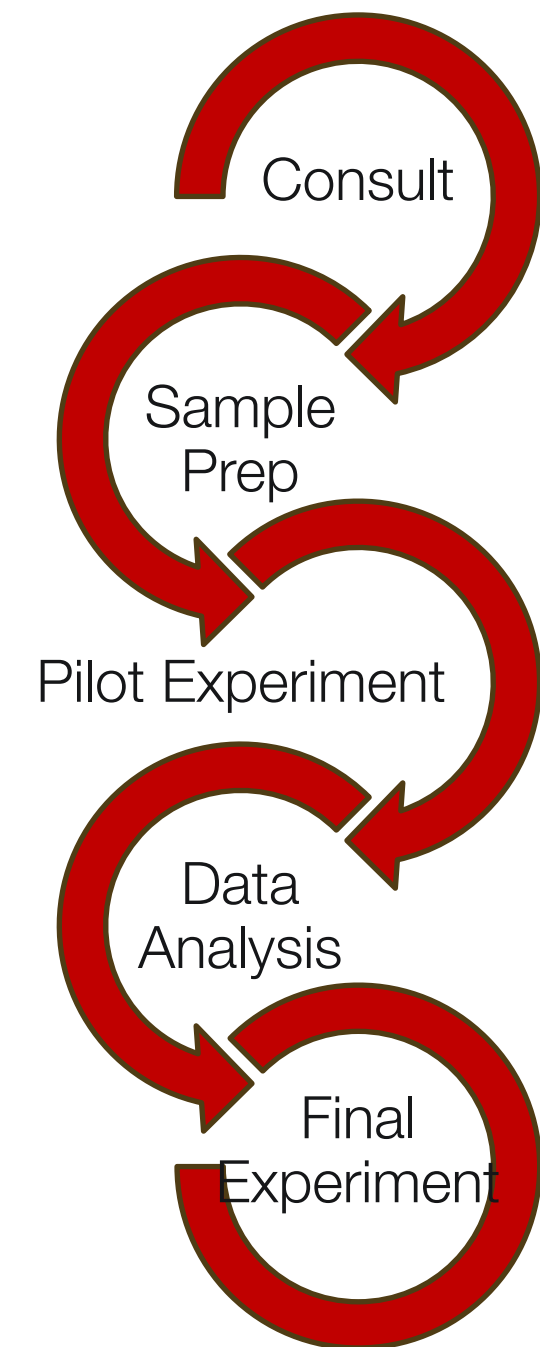
Data Analysis

Final Experiment

STEP I – Pilot experiment planning

What causes technical noise in single cell expts?

"Technical Noise": When non-biological, technical factors cause

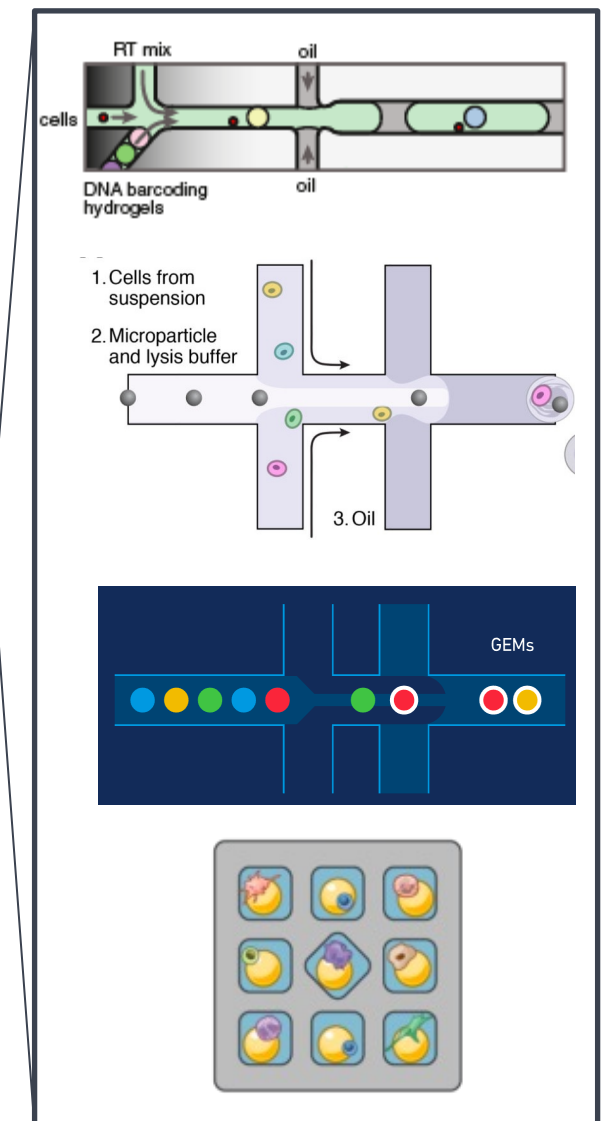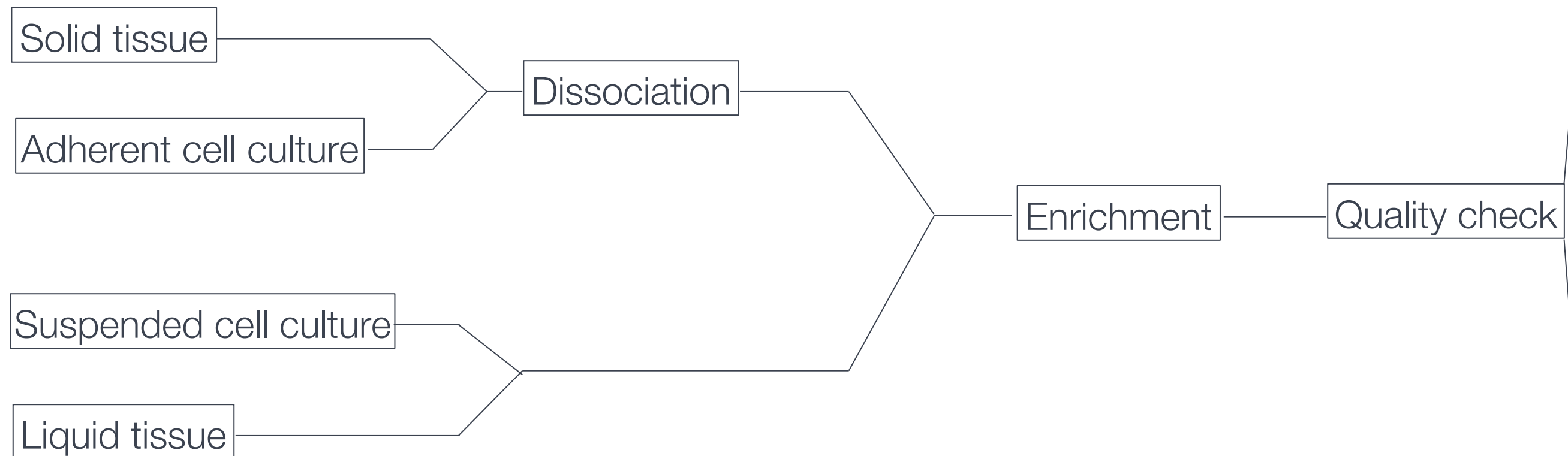changes in the data produced by the expt. leading to wrong conclusions

2 kinds of technical noise -

- Variance resulting from experimental designs and handling (e.g.

  different handling personnel, reagent lots, PCR amp cycles,

  equipment, protocols etc) -> "Batch effect correction"

- Variance resulting from sequencing (e.g. library prep, GC content,

  amp bias etc) -> "Normalization"

Consult

Sample
Prep

Pilot Experiment

Data
Analysis

Final
Experiment

STEP I – Tissue preparation

- What is your sample of interest? How would you obtain that?
    - Which population in a tissue should be examined?
    - What is the abundance of tissue? Does it require enrichment?

STEP I – Tissue preparation

Key considerations for the preparation of a high quality, viable, single-cell suspension

- Dissociation

Table 1: Tissue dissociation protocols

Method	Description	Example protocol/provider
Mechanical	Tissue is mechanically sheared and disrupted through cutting, dicing, pipetting, etc	Isolation of various hematopoietic lineages from bone marrow, spleen, or lymph nodes
Enzymatic	Tissues are incubated with various enzymes such as collagenase, trypsin, dispase, elastase, etc to cleave protein bonds	Worthington Biochemical Corporation
Combinatorial	Mechanical and enzymatic methods can be performed sequentially or simultaneously, with the aid of automated systems, for more extensive dissociation	Miltenyi gentleMACS

Dissociation protocol must be standardized for every tissue – no universal protocol

STEP I – Tissue preparation

Key considerations for the preparation of a high quality, viable, single-cell suspension

- Enrichment (optional but often critical step)

Table 2: Enrichment methods

Method	Description	Available protocol/provider
Centrifugation	Cell populations of interest are enriched based on size, shape, or density by centrifugation through a density gradient medium	Sigma-Aldrich
Bead-based enrichment	Cell populations of interest (including live cells) are enriched by positive/negative selection with magnetic bead–conjugated antibodies	Miltenyi Biotec
FACS	Cell populations of interest (including live cells) are enriched by positive/negative selection with fluorophores/fluorochrome-conjugated antibodies	Beckman Coulter Becton Dickinson BioLegend Bio-Rad

STEP I – Tissue preparation

Key considerations for the preparation of a high quality, viable, single-cell suspension
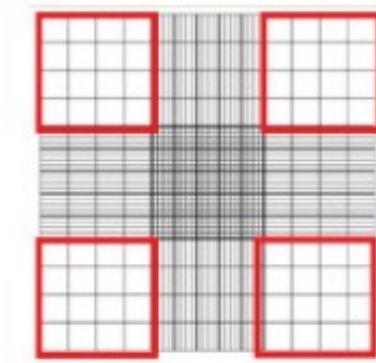
- Enrichment

Table 3: Reagents for separating live and dead cells

Reagent	Mechanism	Pros	Cons
Classic DNA dyes	Membrane impermeant dyes (eg, PI, 7-AAD) that bind DNA will be excluded by live cells	Inexpensive, easy to use	Not compatible with intracellular staining
Amine dyes	Membrane impermeant dyes that bind amine groups of proteins will be excluded by live cells	Compatible with intracellular staining, wide selection of dyes available	More expensive than other dyes, labeling must be done in absence of free protein
Vital dyes	Membrane permeable dye that becomes fluorescent only when cleaved by metabolically active (live) cells	Inexpensive, easy to use	Challenging to use with intracellular staining

The method chosen is driven by various factors

Tissue Preparation: cell numbers

- High thruput platforms need a 10,000-25,000 cell minimum

  - 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

  - Automated cell counters can also give faulty counts

Types of Stains: Acridine orange (marks live cells green)/ Propidium iodide
(marks dead cells red), Trypan Blue (marks dead cells blue)

Tissue Preparation: viability checks

- Check viability of sample over 30-90mins on ice (90-95% viability)

  - If viability decreases over a short period of time this will be reflected in transcriptional data

  - Will see high mitochondrial read counts

- Number of dead cells ∝ number of wasted sequencing reads ∝ wasted $$$

  - 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

  Check single cell suspension supernatant for the presence of "ambient RNA" or free-floating RNA (Ribogreen binds RNA in sol to produce green florescence)

  - Creates background noise in all samples and complicates analysis

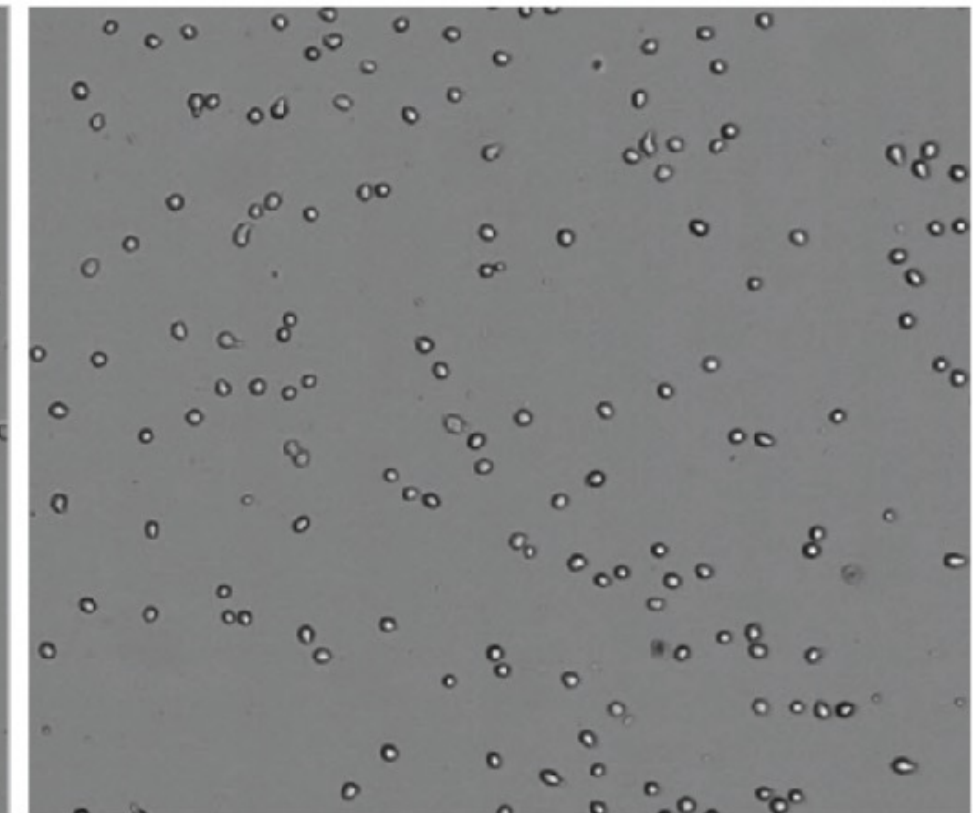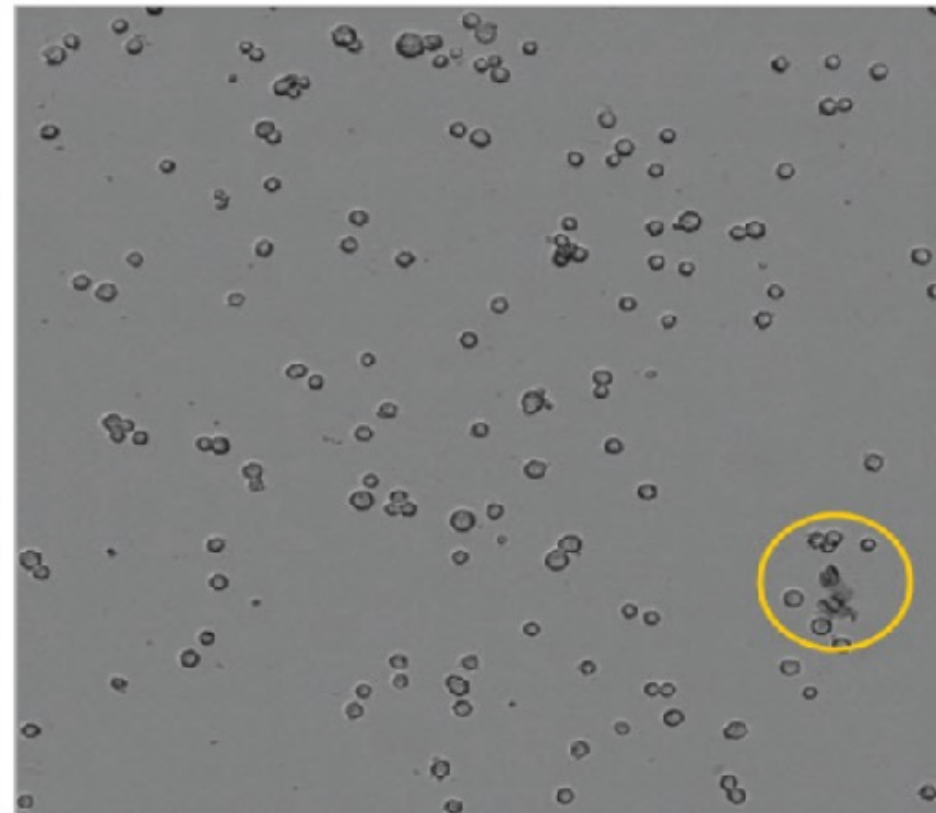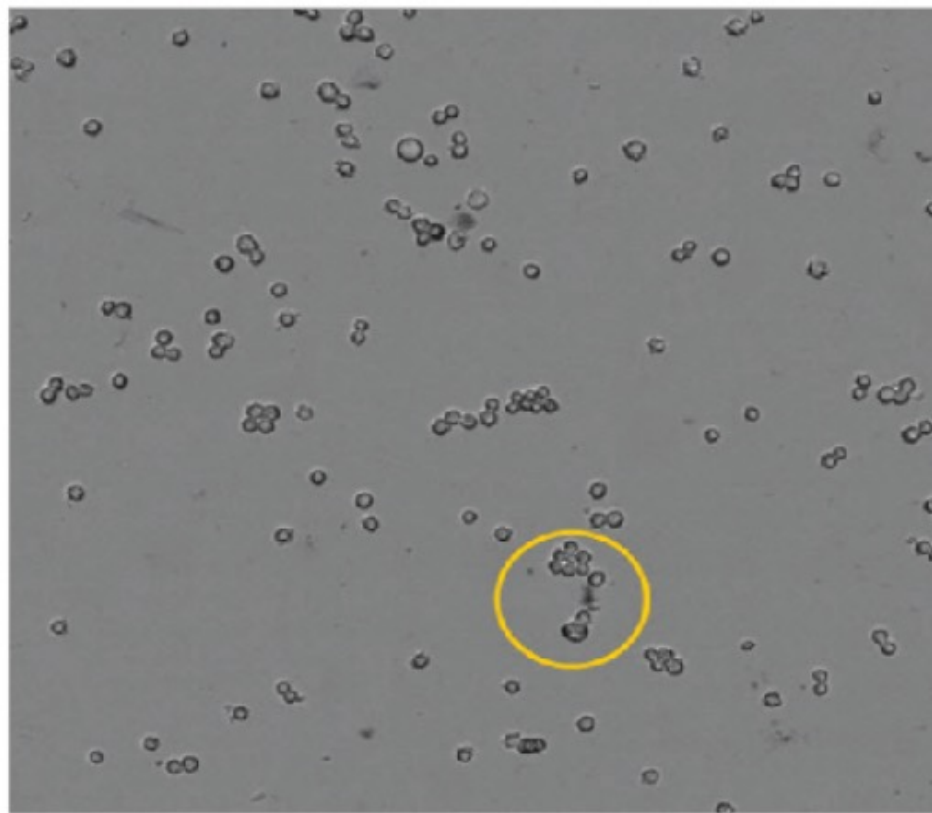# Tissue Preparation: increasing cell viability

- Mild dissociation reagent (TrypLE, StemPro, Accutase, Liberase)

- Shorten dissociation time, reduce dissociation temp

- Adding a DNAse step can help reduce clumping of dead cells.

- Using ROCK inhibitor/ apoptosis inhibitor (esp. epithelial cells)

- Avoid harsh handling: cell pelleting, centrifugation, pipetting etc

- Avoid FACS sorting on fragile cell types (try MACS instead)

- Using correct media: PBS w/ 0.1-1% BSA, 2% FBS in defined media - make sure final buffer does **not** contain calcium, EDTA, or heparin (inhibit RT)

Tissue Preparation: increasing cell viability

- Mild dissociation reagent (TrypLE, StemPro, ~~Accutase, Liberase~~)

- Shorten dissociation time, reduce ~~dissociation time~~

- Adding a DNAse step can help reduce ~~clumping~~ of dead cells.

- Using ROCK inhibitor/ apoptosis inhibitor (esp. epithelial cells)

- Avoid harsh handling, cell pelleting, centrifugation, pipetting etc

- Avoid ~~FACS~~ sorting on fragile cell types (try MACS instead)

- Using correct media: PBS w/ 0.1-1% BSA, 2% FBS in defined media - make sure final buffer does **not** contain calcium, EDTA, or heparin (inhibit RT)

PRACTICE PRACTICE PRACTICE!
This is why the actual run day should not be the first time you attempt entire protocol!

Tissue Preparation – Quality control (QC)

- Visual Inspection

Debris, cell duplets, cell aggregates ✗   < 10% doublets ✓

- Use of flow cytometry to assess multiple metrics simultaneously (viability, size distribution, cell concentration)

Tissue Preparation: How many cells to barcode

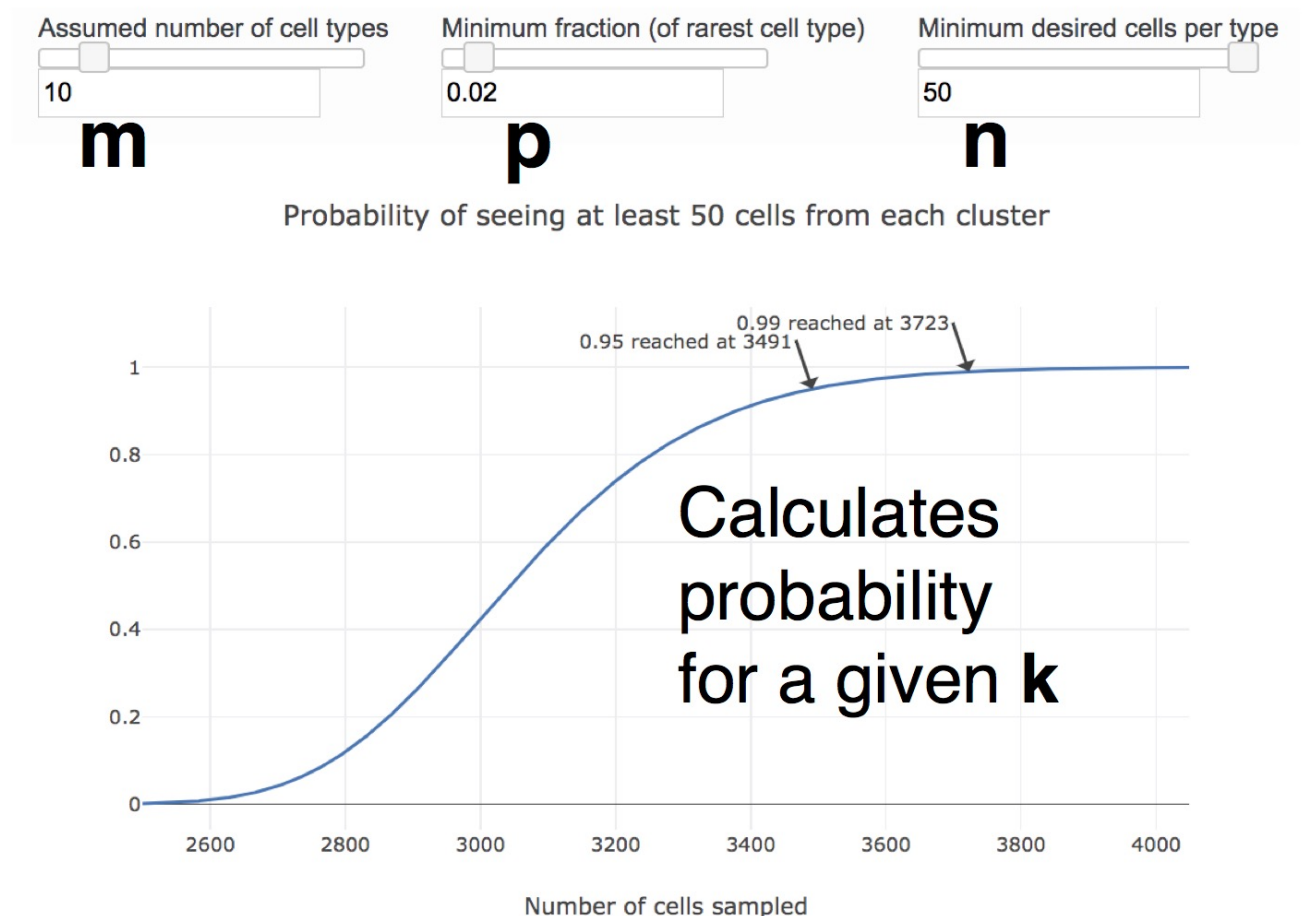How many cells should I barcode or sequence?

- How many high-quality, viable cells do I get from my tissue?

- **What is the biological Q**? What is the % of the smallest sub-population of interest?

(A minimum of 50-100 cells w/ unique transcriptomic signature needed to form

distinct clusters in tSNE/Umap plots)

- Do I need more cells or more sequencing depth?

(More the rare subtypes present, more cells need to be profiled and sequenced)

More you sequence, the higher the cost of your experiment!

Assumed number of cell types: 10 → **m**

Minimum fraction (of rarest cell type): 0.02 → **p**

Minimum desired cells per type: 50 → **n**

Probability of seeing at least 50 cells from each cluster

0.99 reached at 3723
0.95 reached at 3491

Calculates probability for a given **k**

Number of cells sampled

Tissue Preparation: cryopreservation

- Several sc-papers have come out using various cryopreservation techniques on samples

- Success of cryopreservation is dependent on the sample type (e.g. blood and immune cells do great!)

- <mark>Cell viability upon thaw is key to success</mark>

The quality of the tissue at the time of freezing is a major factor in the quality of data downstream
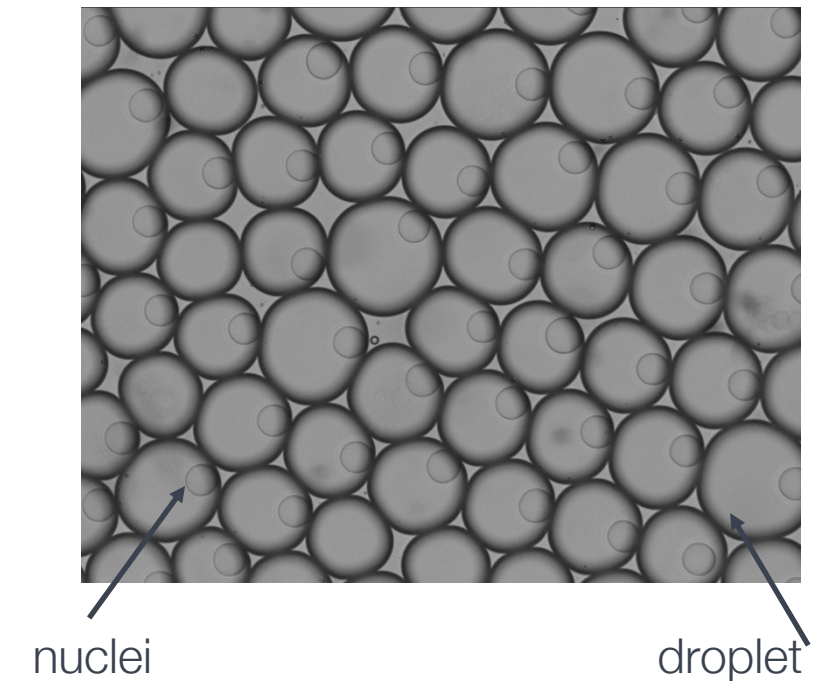
- **Disadv:** you don't know ahead of time if one of your cell types in more sensitive to thawing/death at rehydration, meaning you could heavily bias your sc data if you are not careful!

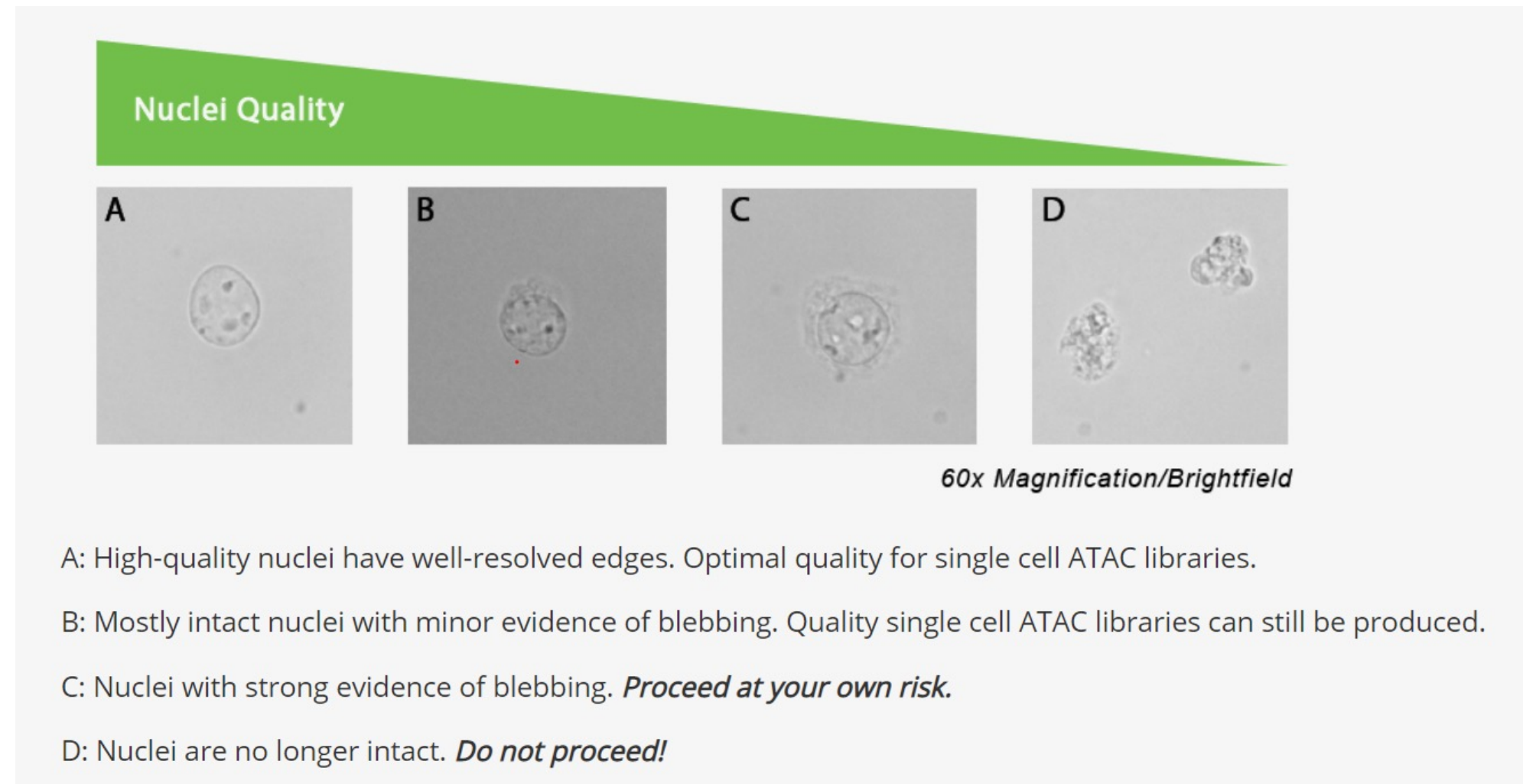Use Std growth media+FBS/DMSO for best results

Tissue Preparation: single nuclei RNA-seq (snRNAseq)

- Extract nuclei from sample of interest

- Removes transcriptional noise from dead/dying cells

- snRNAseq most often used for

  ✓ difficult to isolate/dissociate samples e.g. neuronal samples

  ✓ low viability samples e.g. good for flash frozen clinical samples

  ✓ tissues problematic for sc-processing e.g. adipose tissue, where fat inhibits RT enz in whole cell scRNAseq expt, or pancreatic tissue (high in RNAses)

  ✓ Cell types hard to get from single cell preparations

- Data from scRNAseq is comparable to data from snRNAseq

  ✓ Analysis for snRNAseq different due to presence of introns and non-coding RNA
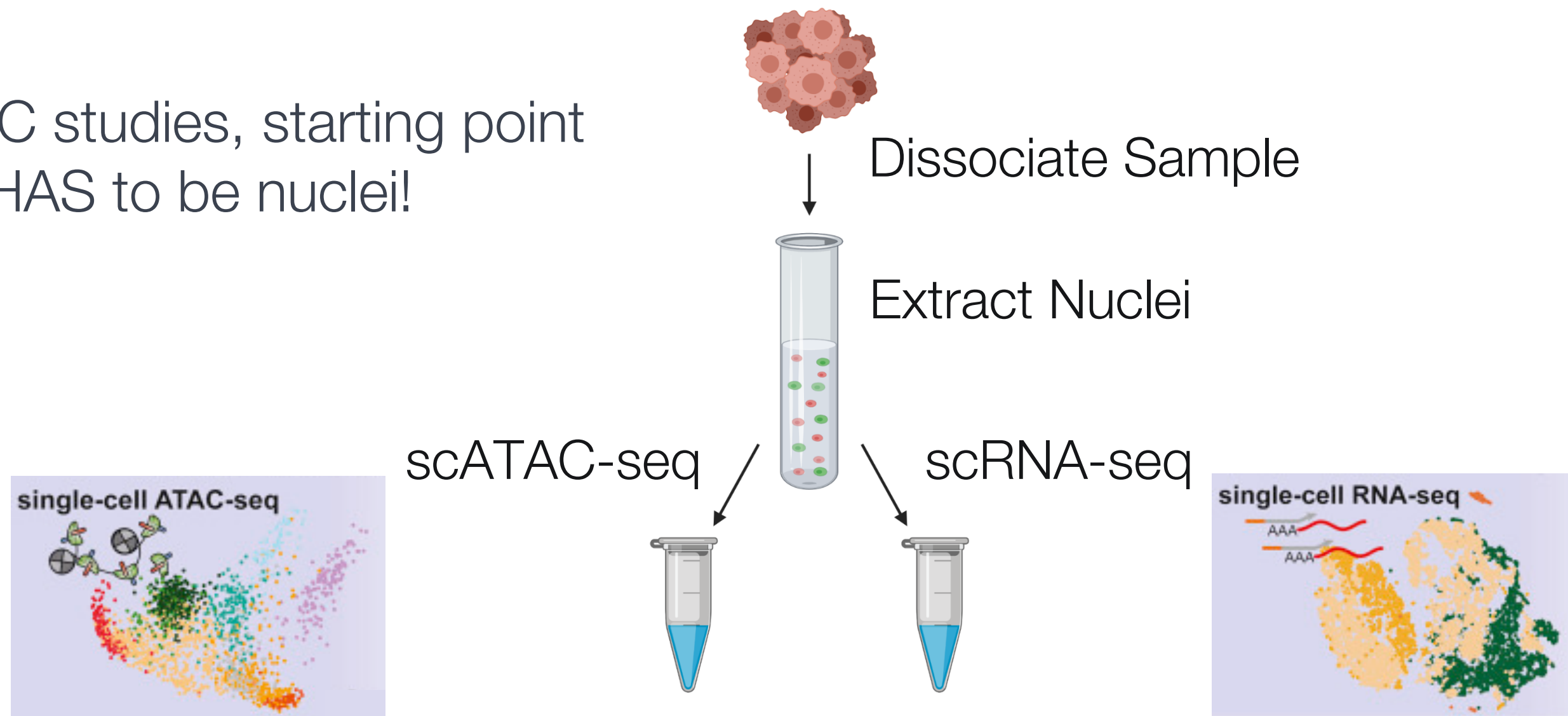

nuclei   droplet

Tissue Preparation: single nuclei RNA-seq

- Good single nuclei suspension. No clumps and minimal debris

- Nuclear membrane integrity is required until cells are encapsulated

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

Tissue Preparation: single nuclei RNA-seq

For ATAC studies, starting point
HAS to be nuclei!

Dissociate Sample

Extract Nuclei

scATAC-seq

scRNA-seq

single-cell ATAC-seq

single-cell RNA-seq

Summary: Best practices to get high quality sample

- Optimize a dissociation protocol that is best-suited for you – no universal protocol for all tissue types!
- Short sample prep time w/ gentle treatments
  - 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 sc-suspension buffer (remove reagents that inhibit RT enz)
- Include biological replicates
- Perform drug/treatment/model vs control on the same day and randomize the order of samples run on different days
- Library prep is the largest source of batch effect – collect all samples from 1 study together then prep as library

Single Cell Core's diverse sample repertoire!

Representative examples
Platyhelminthes (flatworms)
Arthropoda (insects, crustaceans)
Mollusca (squids)
Echinodermata (starfish, brittle stars)
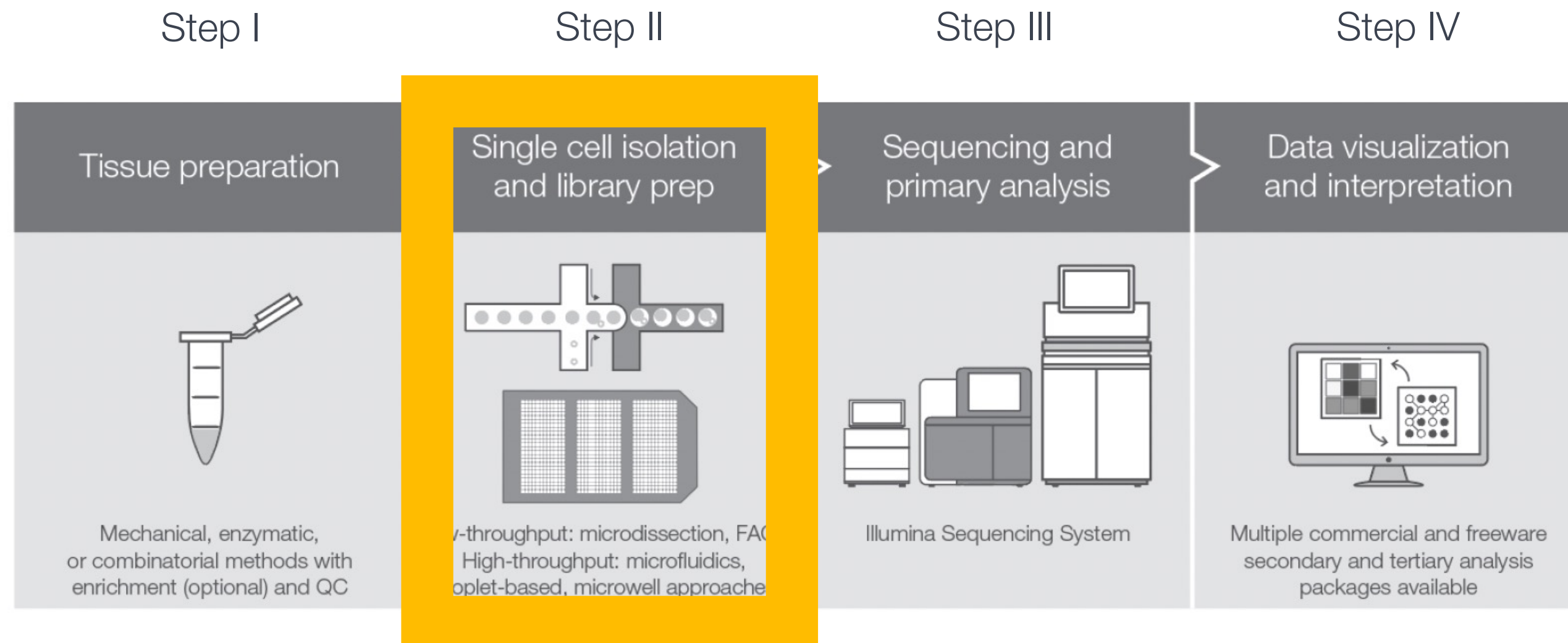Fish (zebrafish)
Amphibians (salamanders)
Aves (chick)
Mammals (human, monkey, mice, rat)

Different Cell lines (human, mice)
& cell types
Organoids (skin, brain, lung, gut)

All major phyla, >50,000 samples

Single Cell Sequencing Workflow – STEP II

Goal: Capture and isolate single cells, prep libraries

STEP II – single cell isolation methods and platforms

Table 4: Low-throughput single-cell isolation approaches

Method	Description	Advantages	Disadvantages	Commercial offering/ Example methods
Serial dilution	Serial dilution of cell suspension down to one cell per well	Simple approach; does not require specialized equipment	Time-consuming, probability of isolating multiple cells	Corning Serial Dilution Protocol
Mouth pipetting	Isolation of single cells with glass pipettes	Simple approach	Technically difficult, random	N/A
Robotic micromanipulation	Isolation of single cells with robotic micropipettes	Positional placement of cells	Requires specialized equipment	An automated system for high-throughput single cell-based breeding. Single cell deposition and patterning with a robotic system.
Laser capture microdissection	Dissection of single cells from tissue sections using a laser	Spatial context is preserved	Technically challenging, potential UV damage to DNA/RNA	Laser capture microdissection of single cells from complex tissues.
FACS	Isolation of microdroplets containing single cells using electric charge	Accurate selection of cell types by size, morphology, internal complexity, and protein expression by antibody labeling	Requires expensive, specialized equipment, cells exposed to high pressure	Beckman Coulter Becton Dickinson Bio-Rad

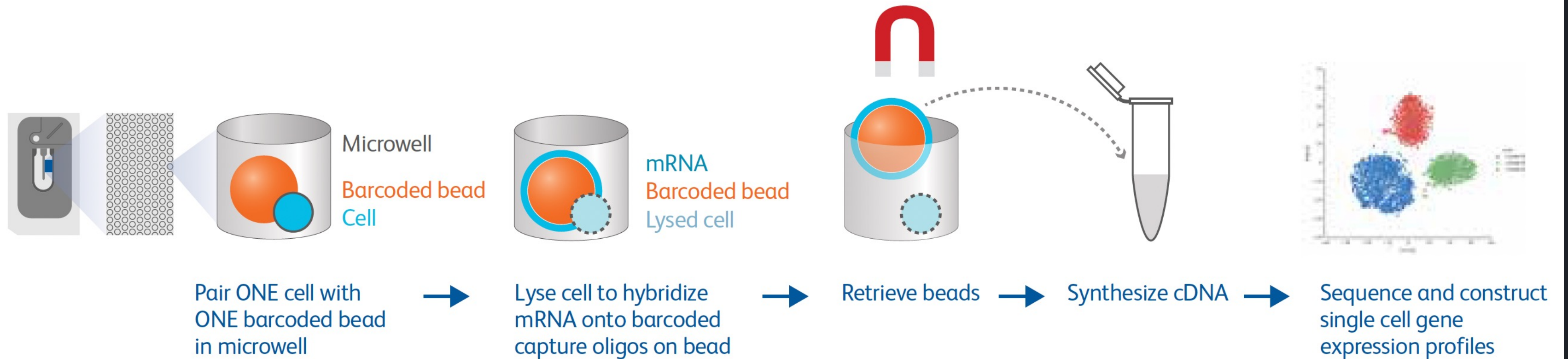Cost effective but sample/cell number prohibitive

STEP II – single cell isolation methods and platforms

Table 5: High-throughput single-cell isolation approaches

Method	Description	Advantages	Disadvantages	Commercial offering/ Example methods
Droplet fluidics platforms[8–11]	Compartmentalization of individual cells in droplets using a microfluidics device followed by lysis and capture of target DNA/RNA	Unique molecular identifiers (UMIs) and cell barcodes enable cell and gene-specific identification, low cost per cell, wide menu of commercial applications	Requires specialized equipment, can be technically challenging	1CellBio inDrop System 10X Genomics Chromium Controller Bio-Rad ddSEQ Single-Cell Isolator Instrument Dolomite Bio Nadia Instrument Mission Bio Tapestri Platform
Microwells[12,13]	Capture of individual cells in microwells of fabricated arrays	Supports imaging and short-term culture of cells, ideal for adherent cells, UMIs enable cell and gene-specific identification	Requires specialized equipment, can be technically challenging	BD Rhapsody Single-Cell Analysis System CellMicrosystems CellRaft AIR System Bio-Rad Celsee Genesis System Takara ICELL 8 cx Single-Cell System

The method chosen will determine lib prep, sequencing and downstream analysis

Microwell based: BD Rhapsody overview

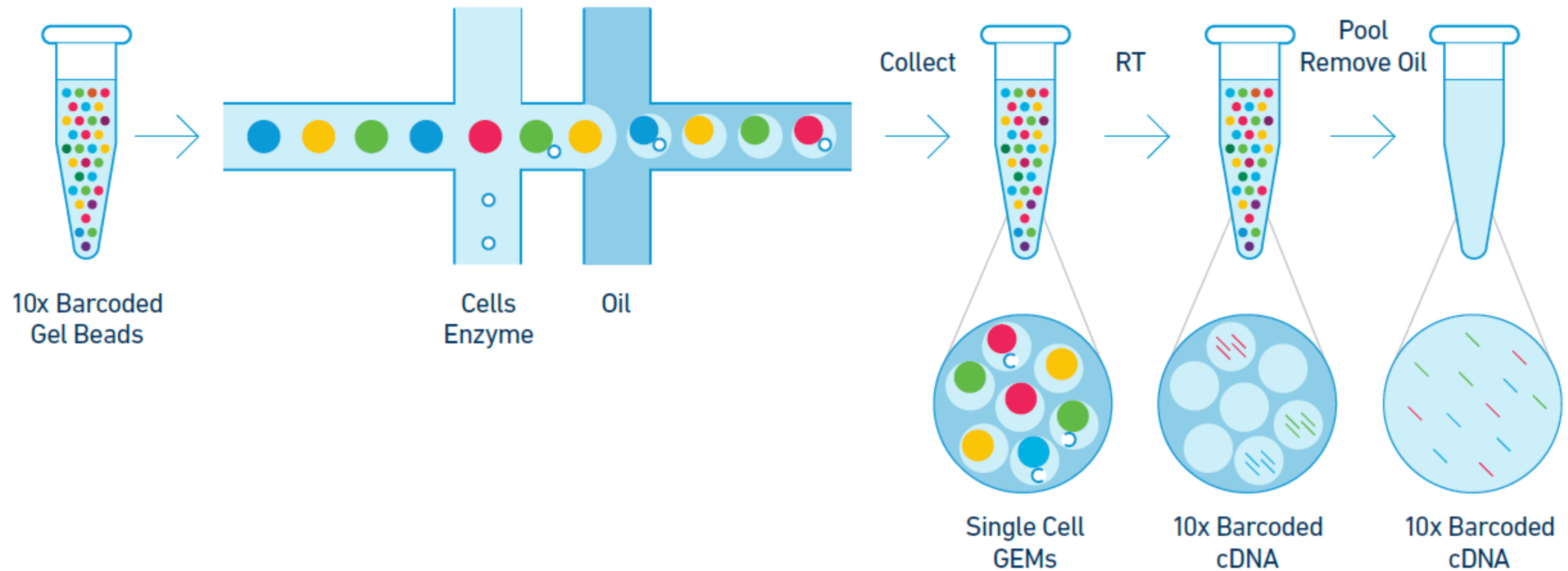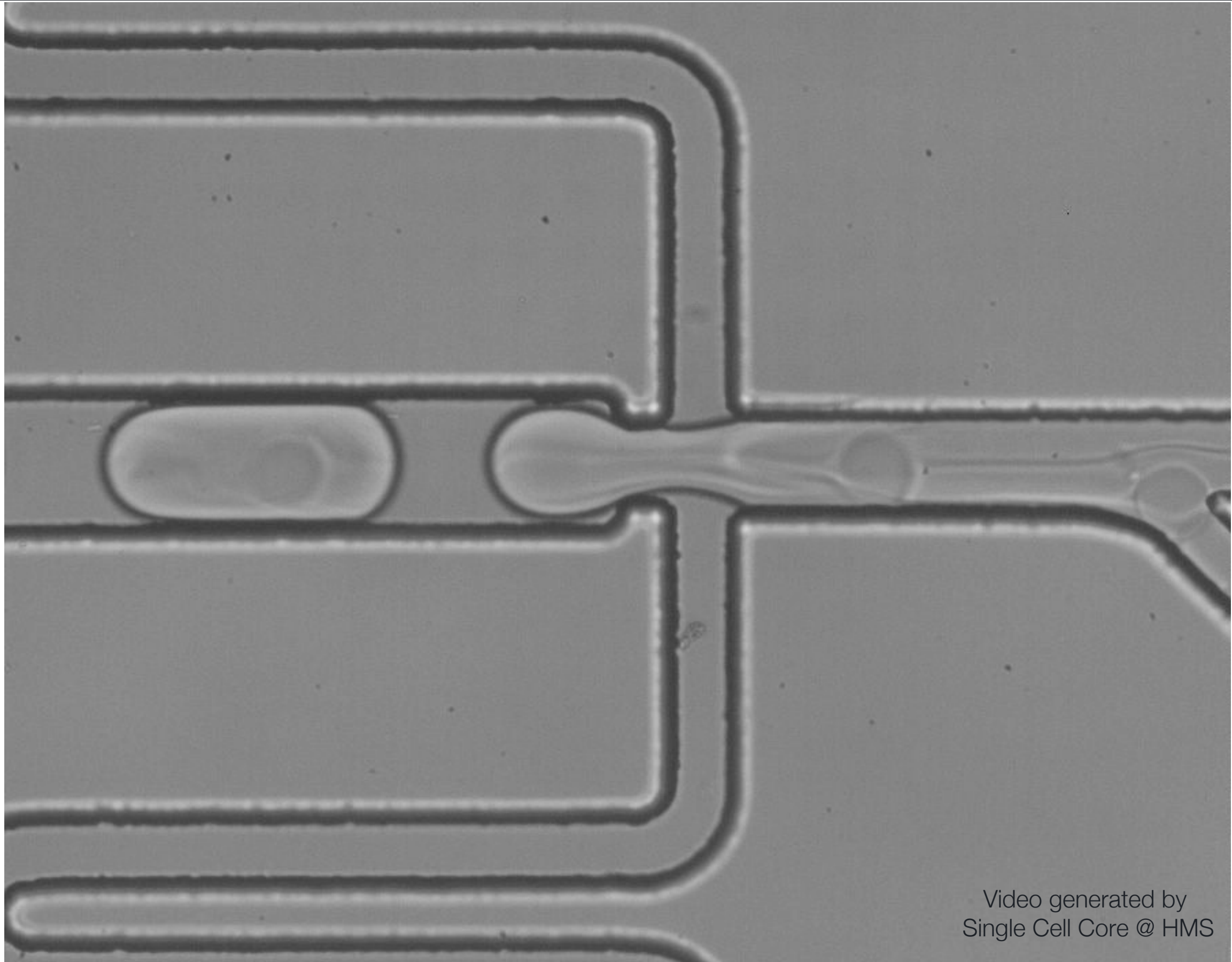| Pair ONE cell with ONE barcoded bead in microwell | → | Lyse cell to hybridize mRNA onto barcoded capture oligos on bead | → | Retrieve beads | → | Synthesize cDNA | → | Sequence and construct single cell gene expression profiles |

Barcoded beads w/ mRNA can be stored long-term, so sample not lost
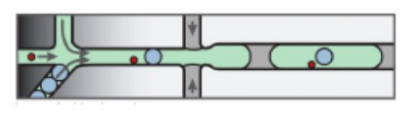
Droplet based: 10x Genomics overview

Droplet based:

10x Barcoded Gel Beads

Cells Enzyme

Oil

Collect

RT

Pool Remove Oil

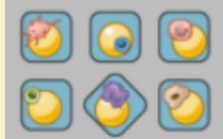Single Cell GEMs

10x Barcoded cDNA

10x Barcoded cDNA

"Gel beads in emulsion or GEMS"

Video generated by
Single Cell Core @ HMS

There are many Single Cell Platforms

	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)

Instrument-free Single Cell Platforms

No complex instrumentation

Highly scalable technology

Flexible configuration

Cost effective

Ideal for pilot Studies

Multi-Omics Capability

Examples -

Fluent BioSci (based on PIPseq)

Parse BioSci (based on combinatorial barcoding)

STEP II: Library prep

Types of information from single cell sequencing –

- Transcriptome – Full length RNAseq (SMARTseq),
  3' or 5' mRNA gene expression libraries,
  Targeted panels or Immune-repertoire V(D)J (T-cell/B-cell receptors)

- Genome – MALBAC, DOP-PCR (whole genome amp from sc)
  Targeted panels

- Epigenome – ATACseq

- Protein capture - CITEseq

Parallel assays to add layered info to scRNAseq data

Multiple libraries from same sample for multimodal sc-analysis
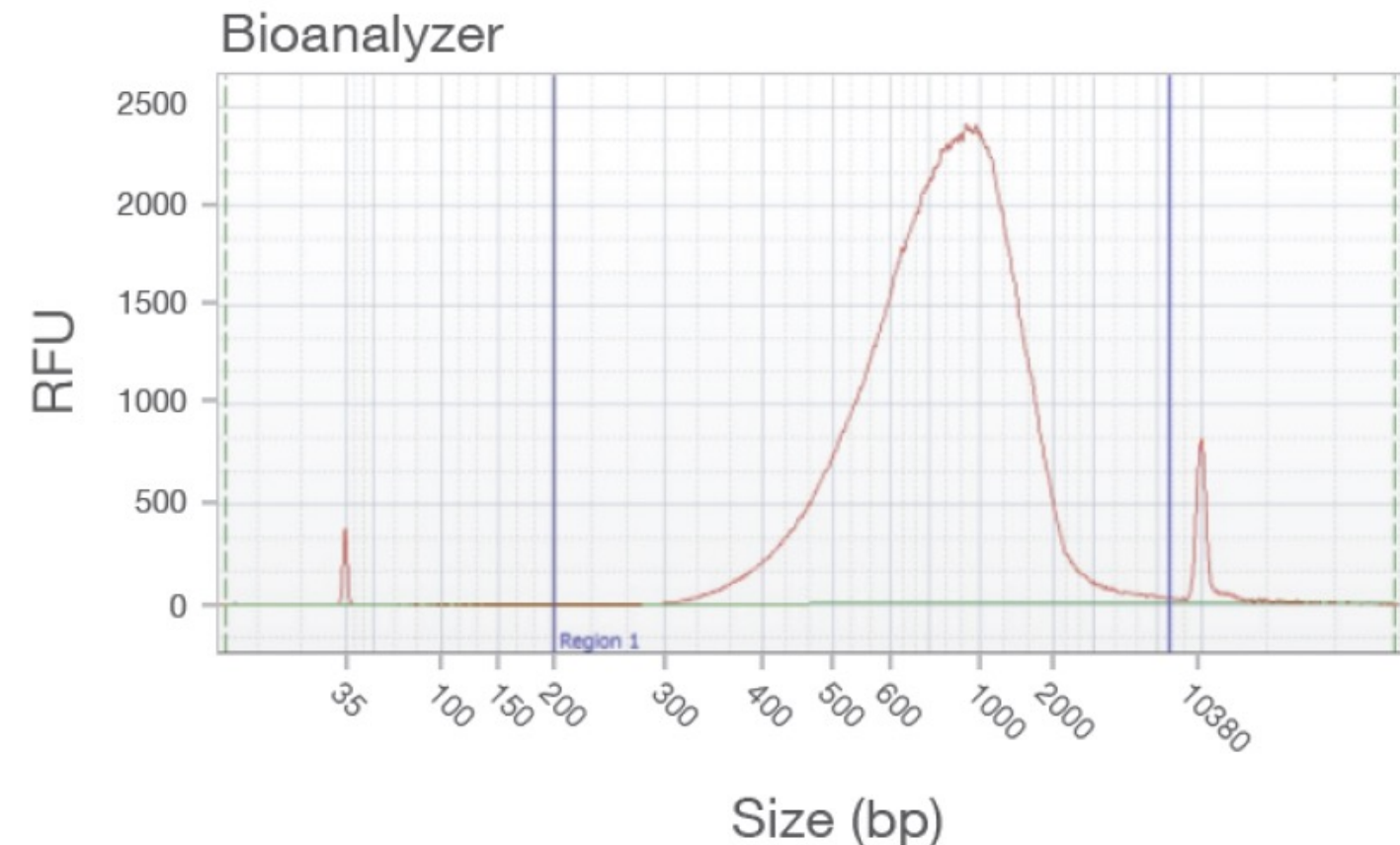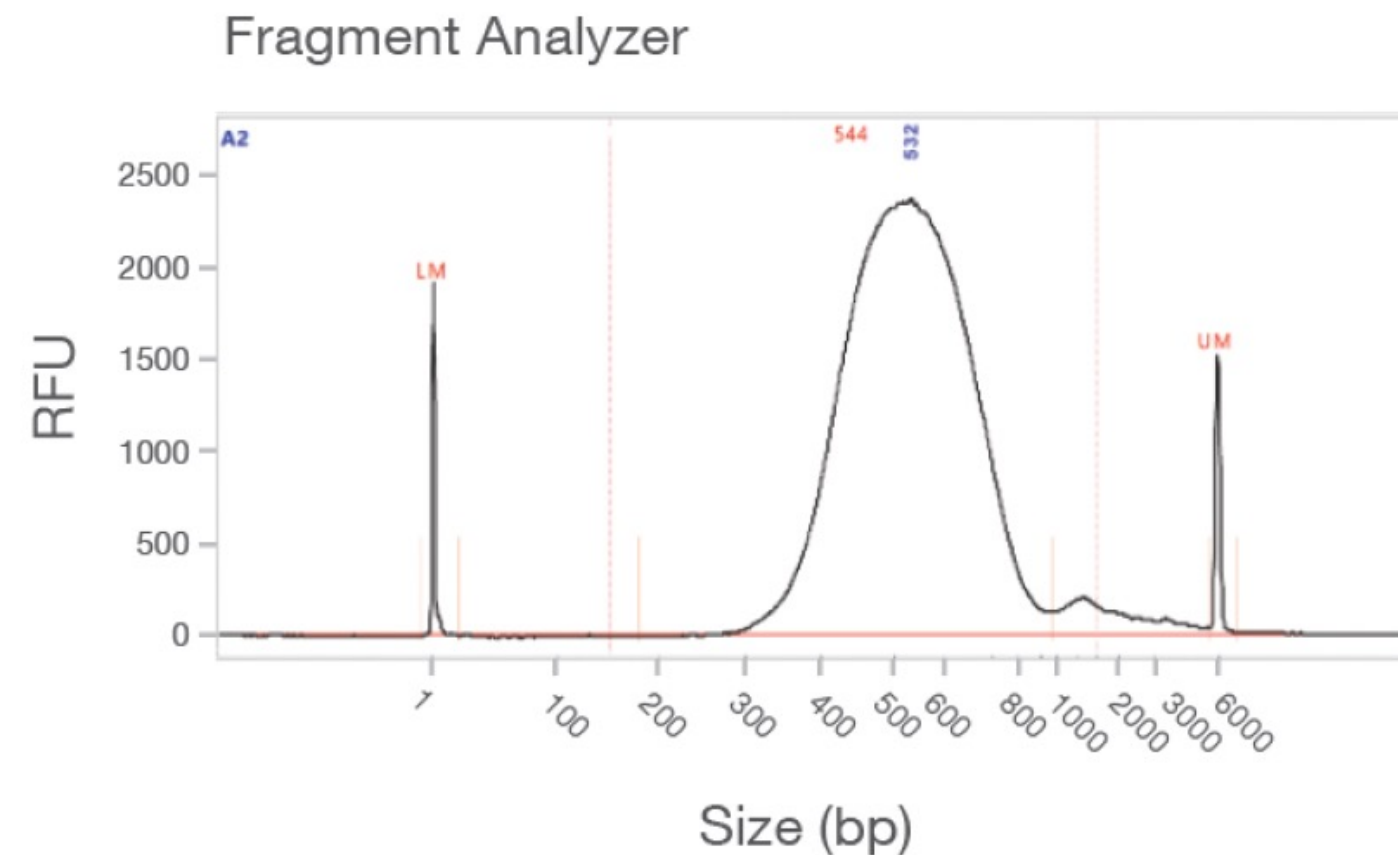
Representative examples:

- scRNAseq (3' or 5' gene expression) + scATACseq (epigenome)

- scRNAseq (3' or 5' gene expression) + CITEseq (surface proteins)

- scRNAseq (3' or 5' gene expression) + cell hashing (surface proteins)

More informative data at same or lower cost!

But expt has to be designed at the beginning for multimodal analysis

STEP II: sc-isolation methods and library prep

- Quality control

Sensitive quantification, size measurements of fragments, detection of possible contaminants

STEP II: sc-isolation methods and library prep

Summary –

- Choose a platform compatible w/ your tissue's biology

- Common platforms – 10x Genomics, inDrops, BD Rhapsody

Note: Not recommended to switch platforms midway through your experiment!

- Library prep – transcriptome, genome, epigenome, or protein capture

- Consider parallel assays to capture maximum info from sc-data at lower costs

- Once high quality lib prepped, proceed w/ sequencing

- Protocol resources – protocols.io, 10x sample prep resources

Single Cell Sequencing Workflow – STEP III

Step I
Step II
Step III
Step IV

Tissue preparation

Single cell isolation and library prep

Sequencing and primary analysis

Data visualization and interpretation

Mechanical, enzymatic, or combinatorial methods with enrichment (optional) and QC

Low-throughput: microdissection, FACS
High-throughput: microfluidics, droplet-based, microwell approaches

Illumina Sequencing System

Multiple commercial and freeware secondary and tertiary analysis packages available

Goal: sequence your sc-lib using a compatible NGS platform

STEP III: Isomolar library pooling for sequencing

- Submitting isomolar libraries for sequencing (equal pooling)

illumına®

Pooling Calculator

Library Plexity			
Unit of Measure for Library	● nM		
	○ ng/µl		
Pooled Library Concentration (nM)			
Total Pooled Library Volume (µl)	150		
Description (optional)			

	Library Concentration (nM)	Library Volume (µl)	10 mM Tris-HCl, pH 8.5 (µl)	Pooling Volume (µl)
Total Pool Volume				

Calculate

STEP III: Sequencing platforms for scRNAseq

Common compatible sequencing systems -

More power/output
Simple benchtop
Affordable

	NextSeq 550 System	NextSeq 1000 and NextSeq 2000 System	NovaSeq 6000 System
Advantages		Power of high-throughput sequencing with the simplicity and affordability of a benchtop system	Unprecedented output and throughput
Ideal for		Mid- to high-throughput sequencing applications and average scale single-cell sequencing studies, such as studies to profile cell function in both development and disease.	Extensive screening studies, such as pharmaceutical screens and cell atlas studies.

STEP III: Considerations for successful sequencing run

Experimental planning - Read depth or 'coverage'

Table 8: Recommended reads for different single-cell sequencing applications

Method	Recommended no. of reads[a]
3' gene expression	15K–50K reads per cell
5' gene expression	50K reads per cell
Antibody sequencing	100 reads per antibody/cell
scATAC-Seq	50K reads per nuclei
5' TCR/BCR	5K reads per cell
Takara SMARTer	1M–2M reads per cell (> 300,000 reads per cell)

The recommended number of reads is based upon manufacturer recommendations

Sequencing depth dependent on sample type and experimental objective

STEP III: Considerations for successful sequencing run

Experimental planning - Read depth or 'coverage'

Example: You have barcoded 10K cells from 4 samples = 40K barcoded cells
40K x 20,000 reads/cell = 800M reads total

NovaSeq 6000 System

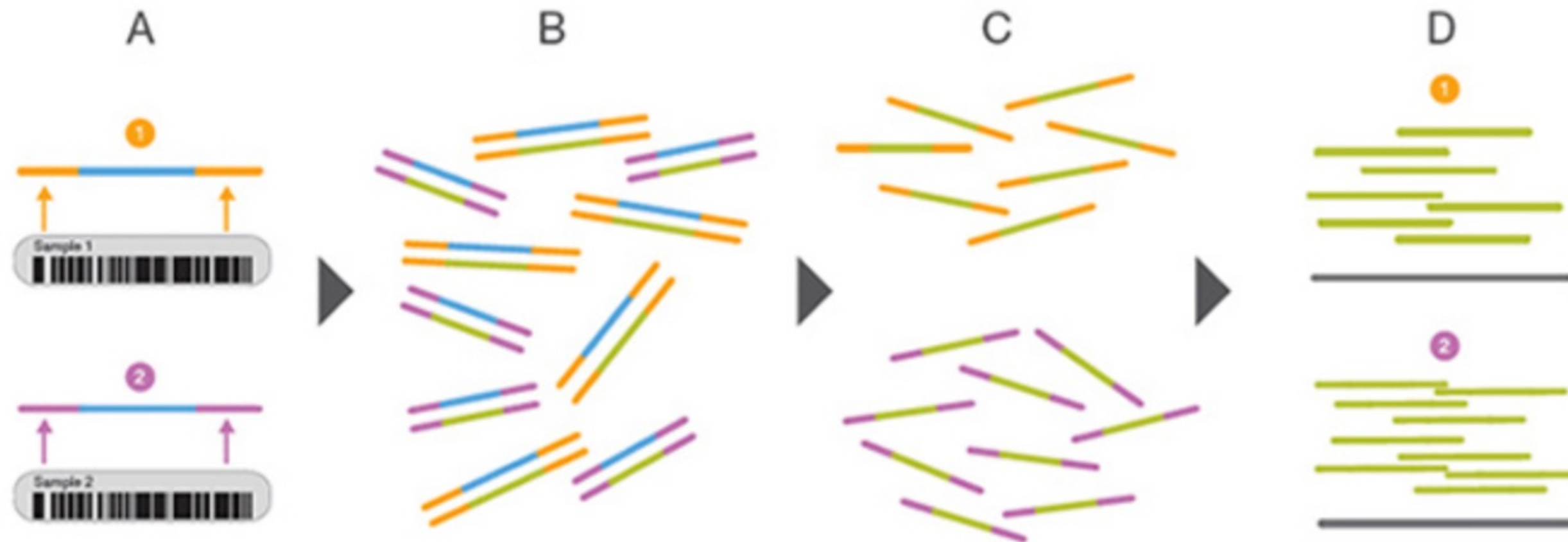| Flow Cell Type	SP ✓	S1	S2	S4
Single-end Reads	650–800 M	1.3–1.6 B	3.3 B–4.1 B	8-10 B
Paired-end Reads	1.3–1.6 B	2.6–3.2 B	6.6–8.2 B	16–20 B

Sequencing jargon

Paired end vs single–read sequencing

Sequencing jargon
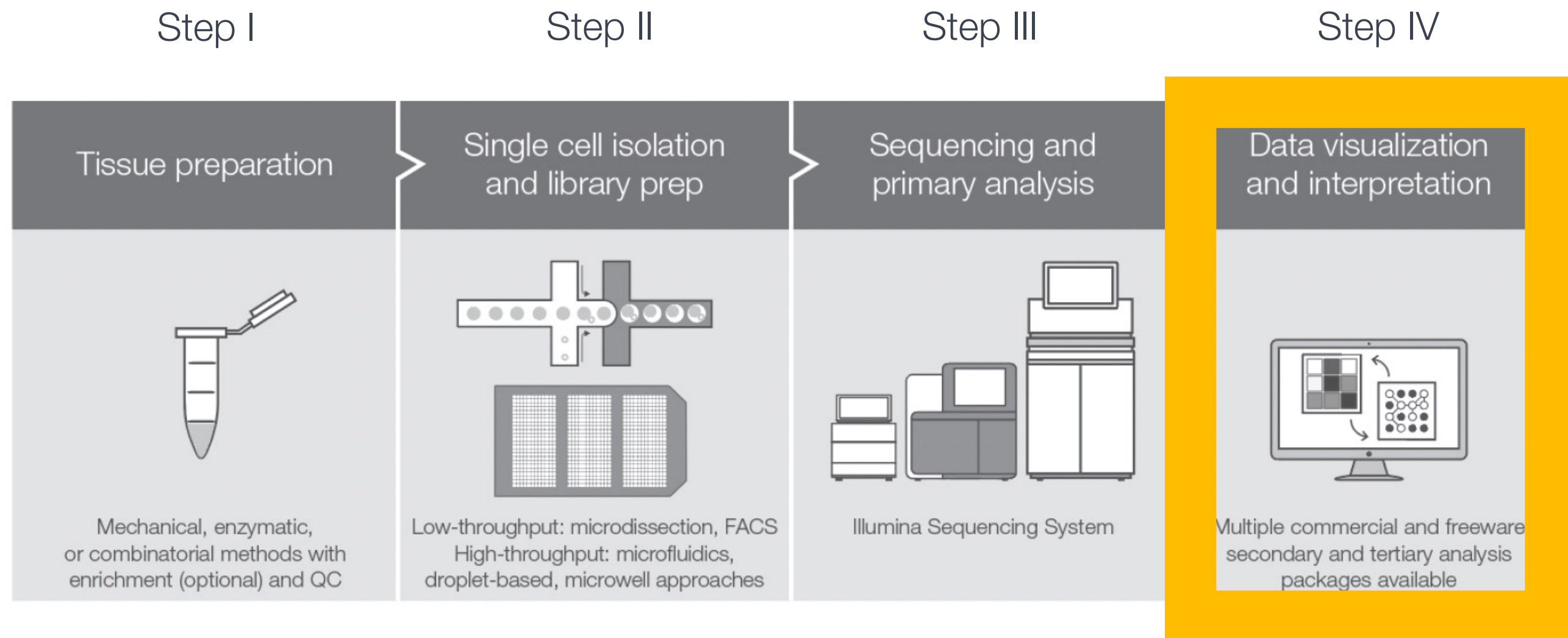
Multiplexing vs demultiplexing

STEP III: sequencing

Summary –

- Choose a sequencing platform compatible w/ your tissue's biology

- Common compatible platforms – NextSeq and NovaSeq
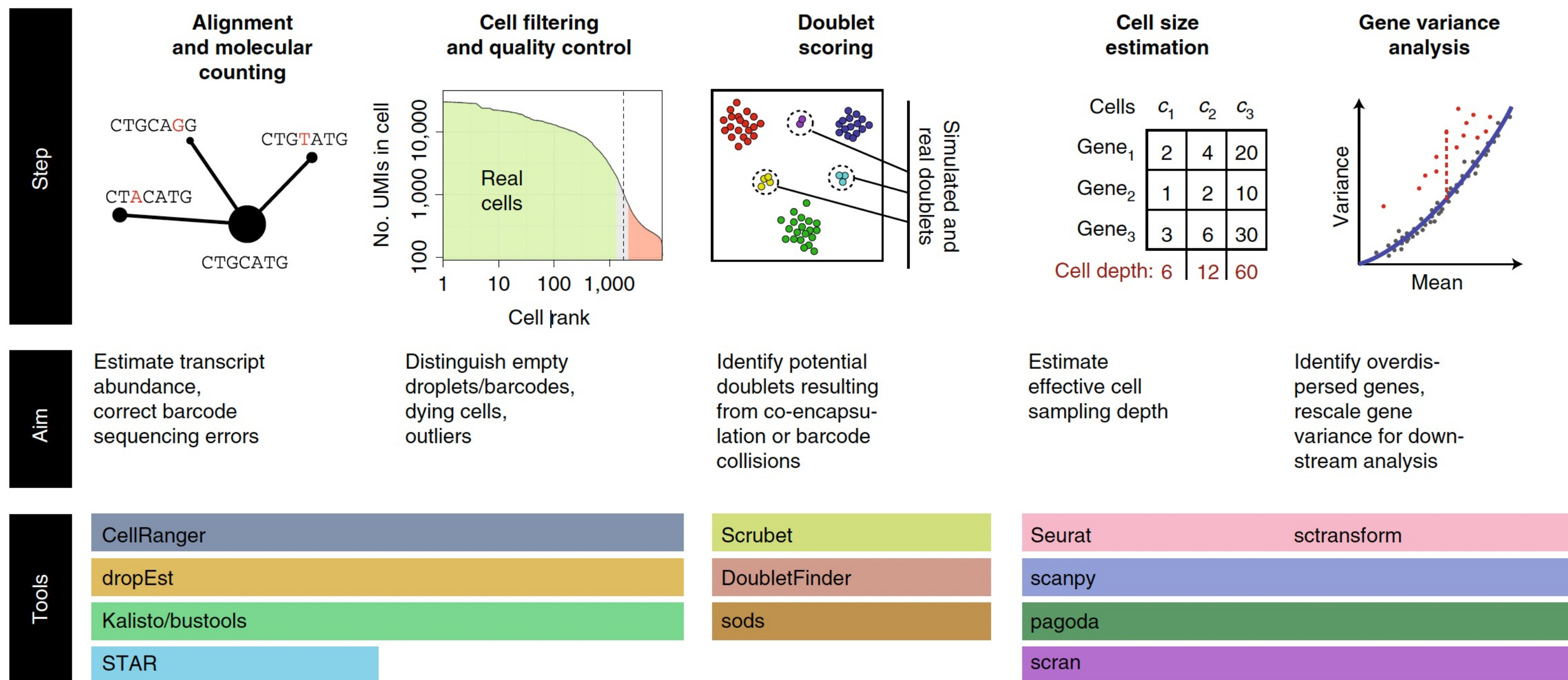
- Go with platform specific recommended sequencing depth

(between 20K-50K reads/cell for GEX)
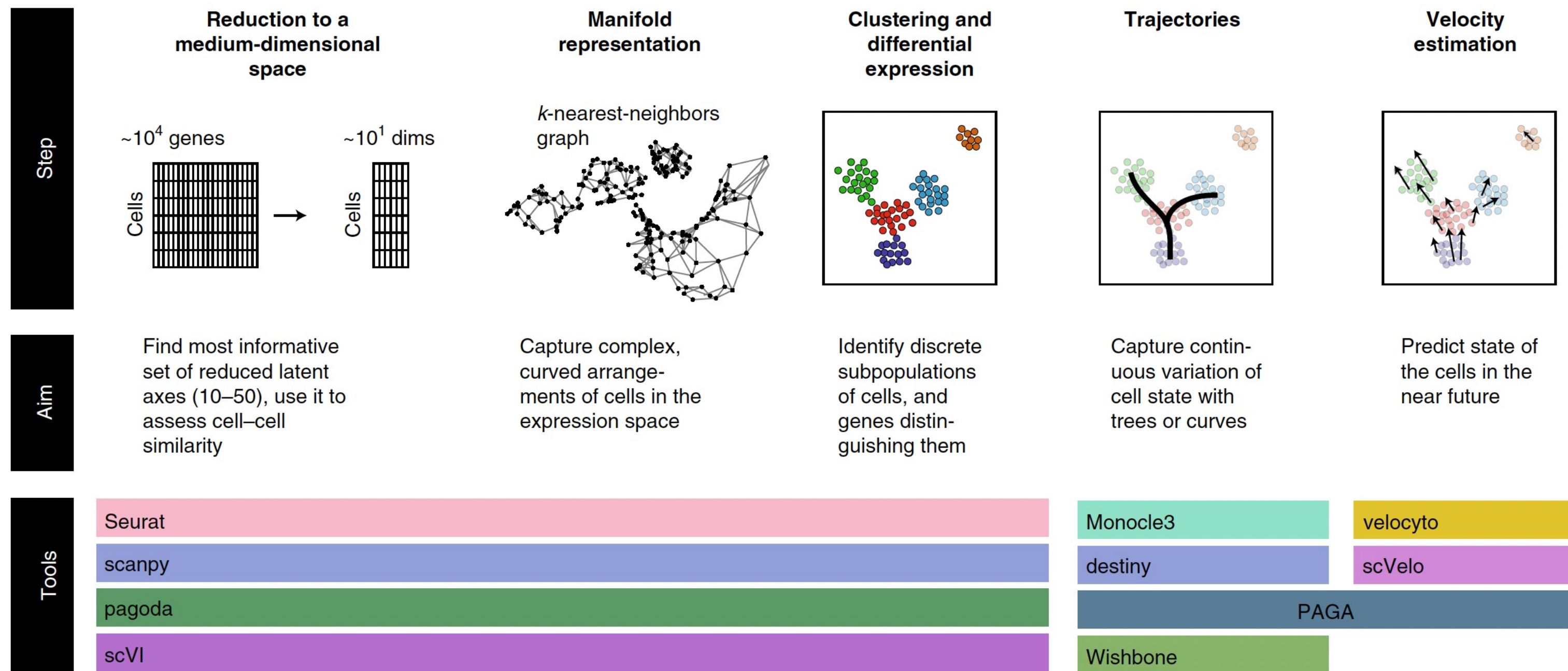
Single Cell Sequencing Workflow – STEP IV

Step I

Step II

Step III

Step IV

Tissue preparation

Single cell isolation and library prep

Sequencing and primary analysis

Data visualization and interpretation

Mechanical, enzymatic, or combinatorial methods with enrichment (optional) and QC

Low-throughput: microdissection, FACS
High-throughput: microfluidics, droplet-based, microwell approaches

Illumina Sequencing System

Multiple commercial and freeware secondary and tertiary analysis packages available

Goal: Analyze, visualize and interpret your sc data

Key preprocessing steps in scRNAseq analysis

Kharcheno, 2021

Key analysis steps in scRNAseq analysis

Kharcheno, 2021

STEP IV: data visualization

Summary –

- Precise pipeline for sc analysis variable and depends on research objectives of study

- Common steps in analysis – primary, secondary and tertiary phases

- In these steps, sequences are aligned, data is characterized, visualized and explored

Now we will learn how to do some scRNAseq analyses!! Enjoy the workshop!

Thank you!

Questions?