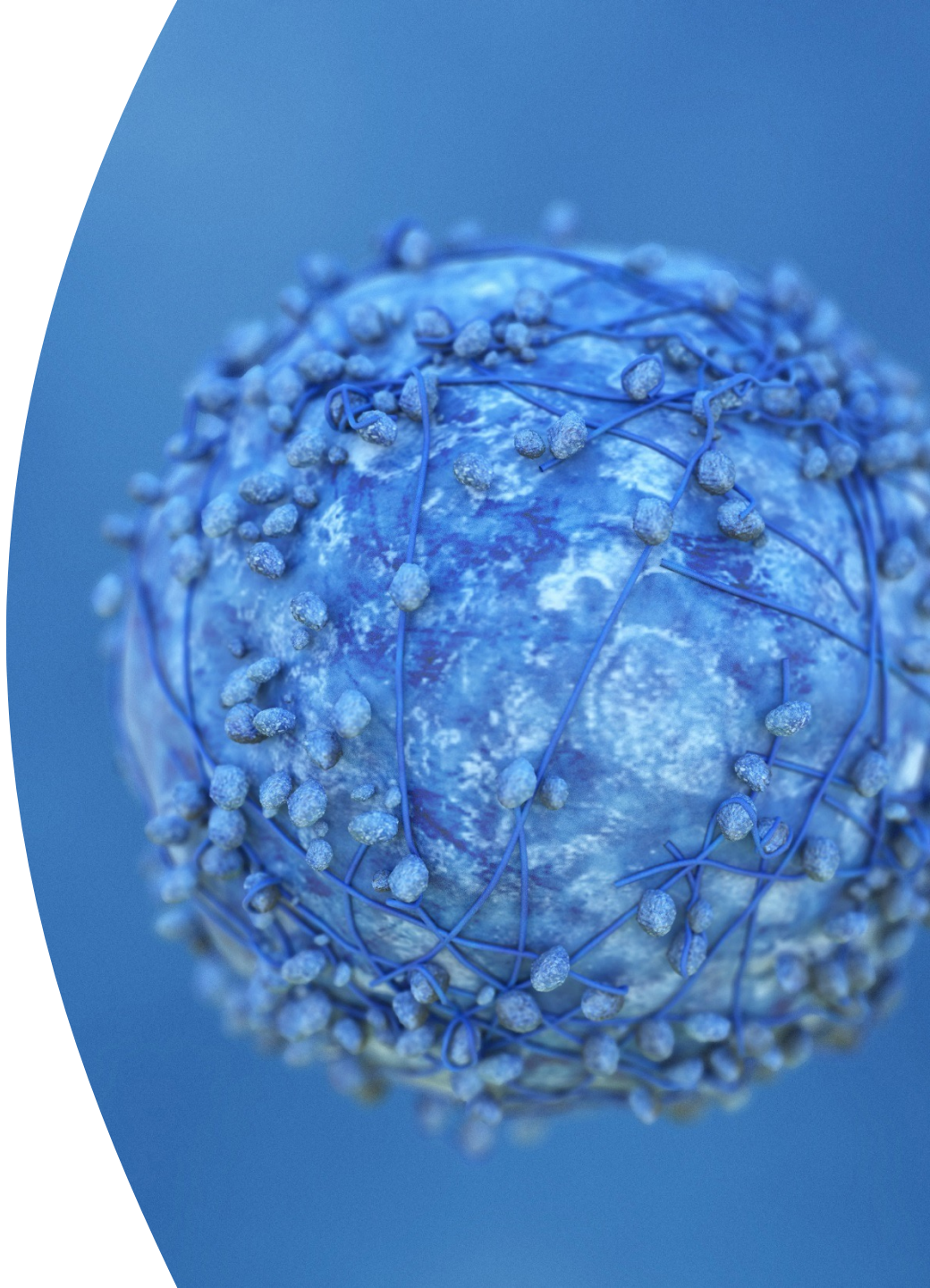
INTRODUCTION TO SINGLE CELL TECHNOLOGIES

Abbi Edwards

Bioinformatics Core



CANCER RESEARCH UK | CAMBRIDGE INSTITUTE

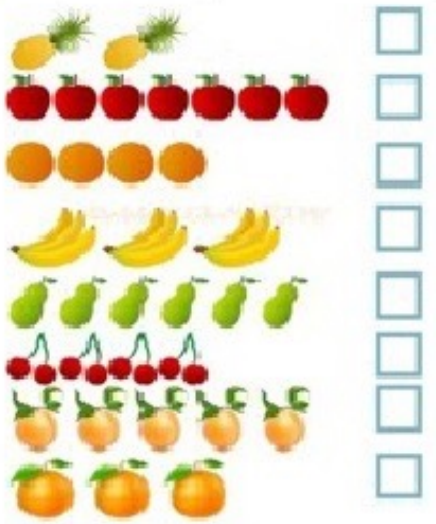
Together we will beat cancer

BULK VS SINGLE CELL

- Average levels
- Comparative transcriptomics
- Disease biomarker detection
- Chromatin Accessibility
- Copy Number and Mutation calling
- Only homogenous systems

Bulk

Single Cell



- All the same techniques but with separate populations
- Define heterogeneity
- Identify rare cell populations
- Cell population dynamics

CANCER RESEARCH UK | CAMBRIDGE INSTITUTE

BULK VS SINGLE CELL RNA-SEQ

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 ~3000 cells / well
   - 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. (even GAPDH)
   - If you only looked at transcripts observed in all cells numbers drop dramatically.

Source: Sarah Boswell, Harvard Medical School, September 2020

TECHNOLOGIES

Droplet Based (10X Chromium)

- Each cell contained in a droplet
- Good For
  - Almost everything
  - Lots of cells
  - Want more than RNA-seq
- Multiplets

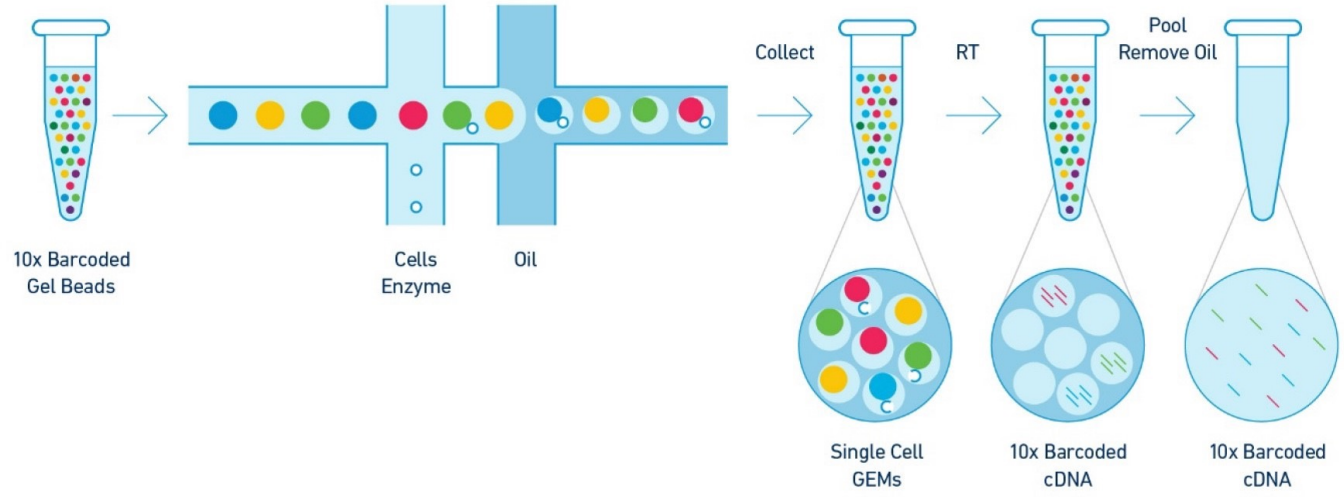


Plate Based

- Each cell in its own tube
- Good For:
  - Low cell numbers
  - Delicate cells
  - Limited material
- SMART-seq (v4)

DISADVANTAGES OF SINGLE CELL?

- Dropouts, noisy data and a lot of potential for bias
- Lowly expressed genes might be undetected
- Samples will contain doublets
- Replicates without batch effect are unlikely
- Expensive
- Difficult to analyse?

WHAT CAN YOU DO?

RNA

- Standard mRNA-seq (Fresh, Fresh Frozen, FFPE)
- Immune repertoire analysis
- CITE-seq
- Crispr Screens and TAP-seq

DNA

- Single cell ATAC-seq (nuclei not cells)
- Mutation calling
- Copy number analysis

Multi-omic Approaches

- ASAP-seq, TEA-seq, BEAM etc.

Hacked methods

- mtsc-ATAC-seq (permeabilised cells not nuclei)

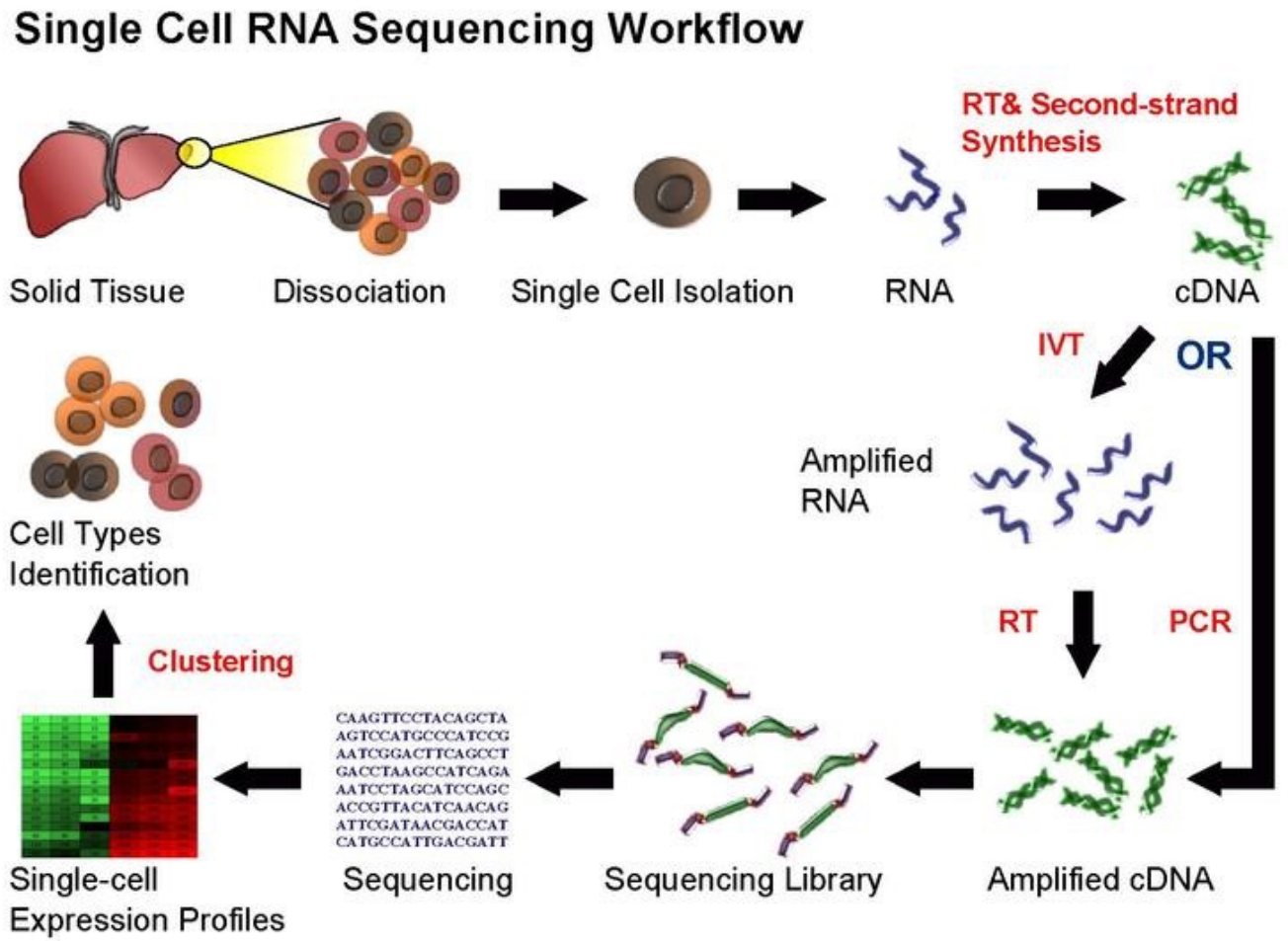
READ TYPES

- Different methods give you different parts of the transcript
- Plate based – full transcript
- Droplet based – either 5’ or 3’ end only depending on chemistry

Full	5’	3’
SMART-seq	GEX	GEX (standard)
Parse	VDJ	Crispr Screen
‘Others’	ATAC-seq	Spatial RNA
	CITE-seq	
	Crispr Screen	

“GOOD SAMPLES”

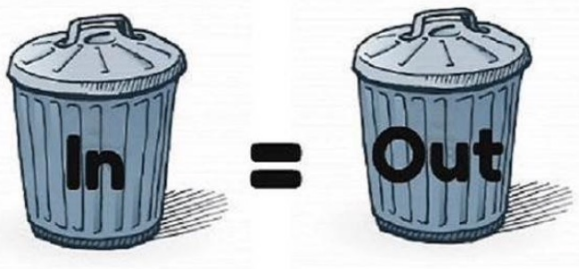
- Most important part to optimise
- Fixing
- Cryopreservation
- Dissociation
- Cell sorting
- Tissue specific buffers



Source: https://en.wikipedia.org/wiki/Single_cell_sequencing

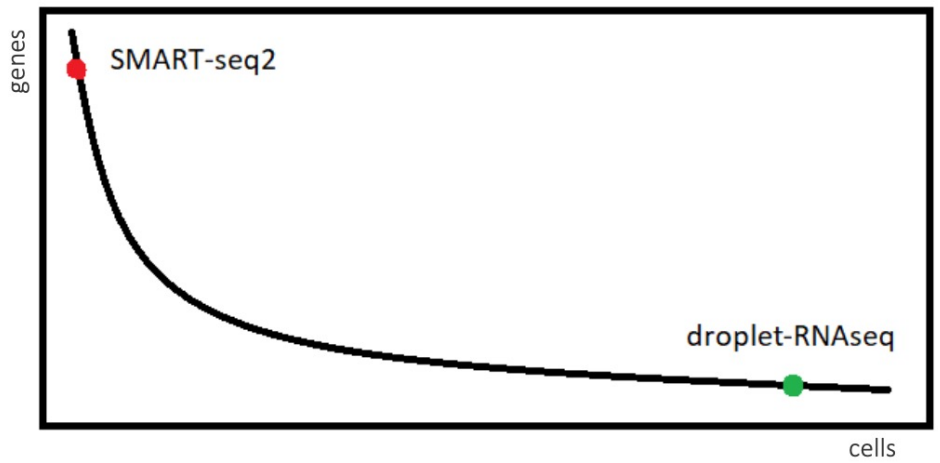
CANCER RESEARCH UK | CAMBRIDGE INSTITUTE

SAMPLE PREPARATION



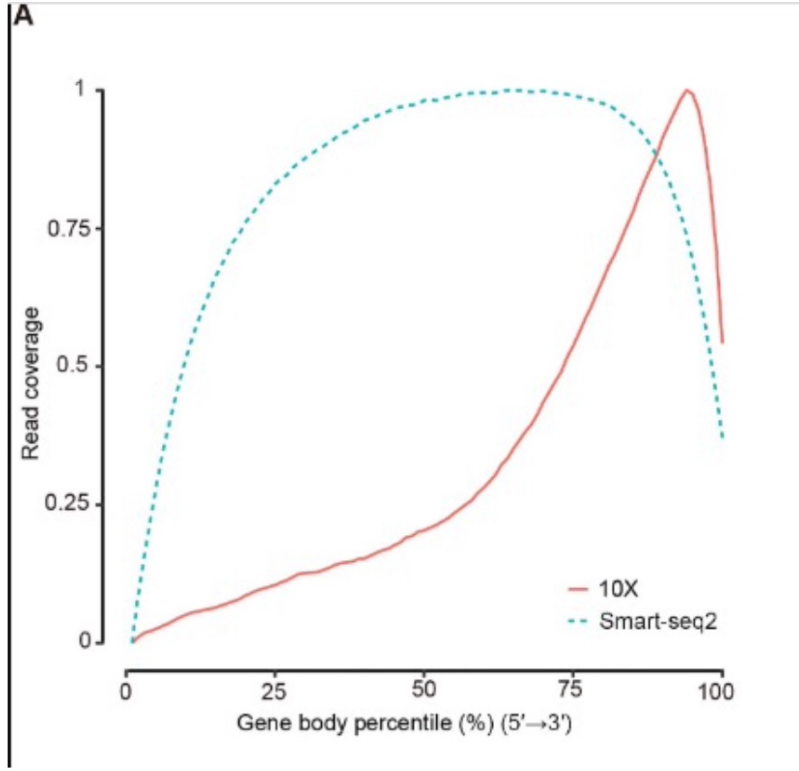
- Understand well the nature of the sample (sampling conditions, preparation, purity)
- Identify the source of technical difficulties in order to resolve them first
- Practice your sample preparation, optimise the protocol well, do not rush to the final experiment
- A well planned pilot experiment is essential for evaluating sample preparation and for understanding the required number of cells.
- You need your cells to be highly viable (>90-95%), have no clumps and no debris. Cell membrane integrity is a must!
- Free-floating RNA will make analysis more challenging
- Be cautious about FACS (especially with more fragile cells). If FACS necessary for enrichment, remember that time is crucial factor
- Count with haemocytometer or cell counter (Countess II Automated Cell Counter) – do not trust sorter counts
- Fixation and cryopreservation are not compatible with many techniques

MORE CELLS OR MORE GENES?



SMART-seq2
- 100 cells
- Full-length libraries
- 1M reads per cell

Droplet-RNAseq
- 10000 cells
- 50k reads per cell
- 3’/5’ bias



Source: Wang, et al. Genom. Proteom. Bioinform. 19(2), 253-266 (2021).

CANCER RESEARCH UK | CAMBRIDGE INSTITUTE

- Required number of cells increases with complexity of the sample.
- Number of reads will depend on biology of sample
- Cell-type classification of a mixed population usually requires lower read depth
- You can always re-sequence your samples.

REPLICATES

- LOTS! £££££££
- Single cells within a sample are not independent of each other
- Using cells as replicates amounts to studying variation inside an individual
- We want to study variation across a population
- Shallow sampling

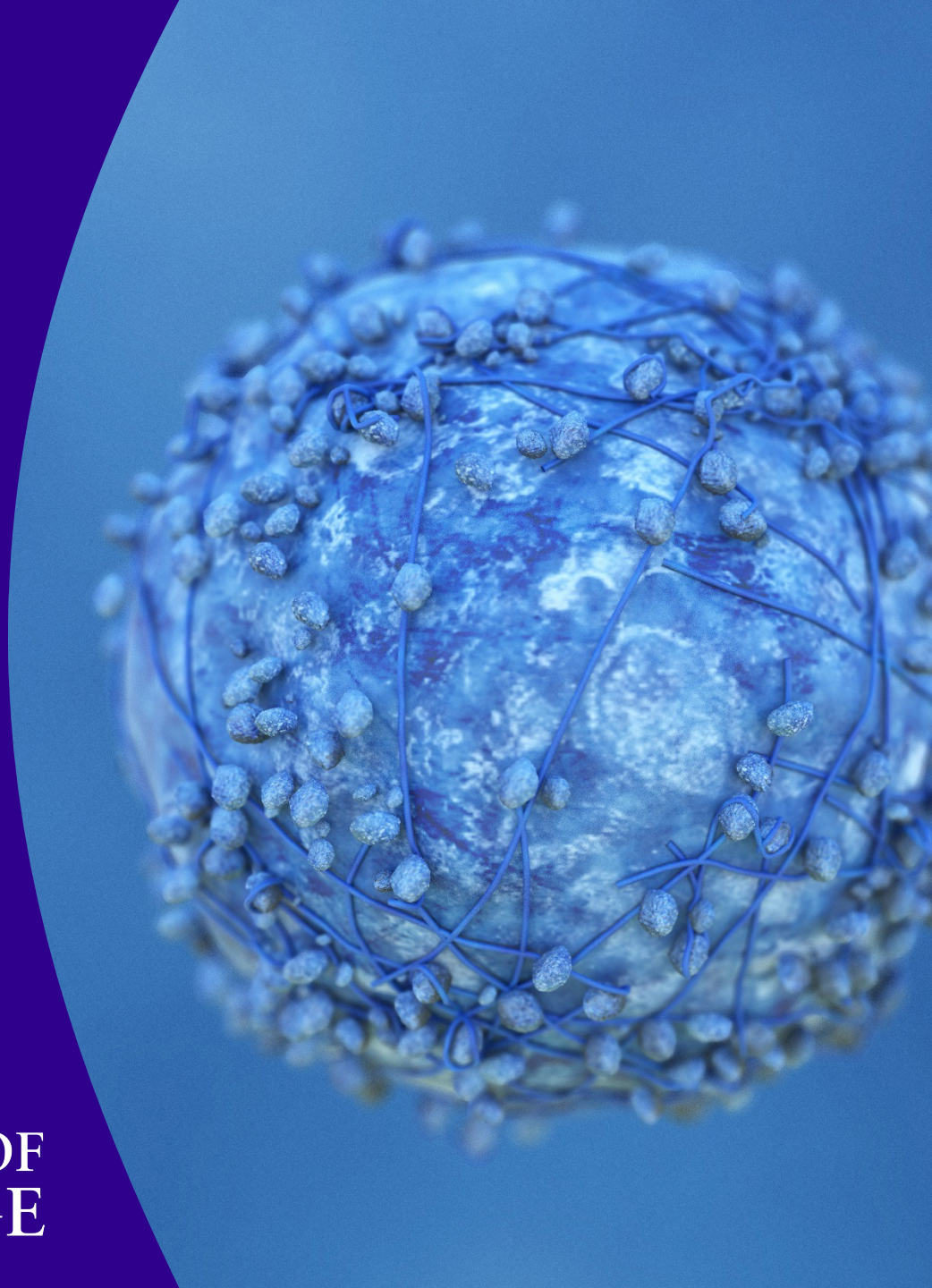
WHAT CAN YOU DO AT THE CI?

(START IN THE GENOMICS CORE)

- All 10X single cell offerings
  - RNA-seq
  - Multiome (RNA + ATAC)
  - Cell surface proteins
  - CRISPR screens
  - Immune profiling
  - Open to collaboration on other related methods
- Spatial (Visium and Xenium or combined)
- Oxford Nanopore
- SMARTseq (plate-based)

(AND WE CAN HELP WITH THE ANALYSIS)

ANY QUESTIONS?

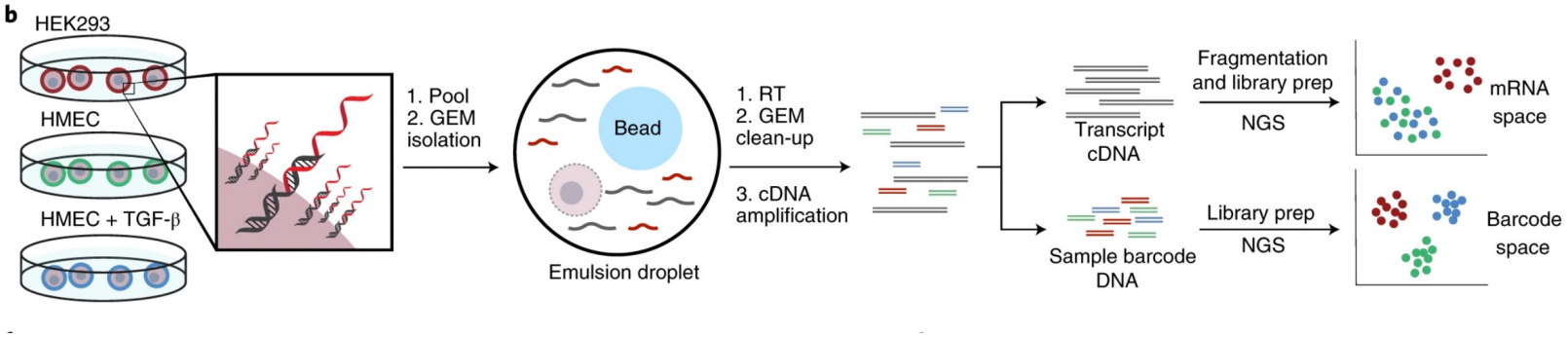


CANCER RESEARCH UK | CAMBRIDGE INSTITUTE

UNIVERSITY OF CAMBRIDGE

CELL HASHING

- Pros: Fits more into a sequencing run
- Cons: Adds another step to the protocol



CANCER RESEARCH UK | CAMBRIDGE INSTITUTE

PROCESSING NON-10X DATA

- SMART-seq can be aligned and counted like bulk RNA-seq data
- Parse biosciences has a cellranger-like piece of software and downstream processing similar to 10X