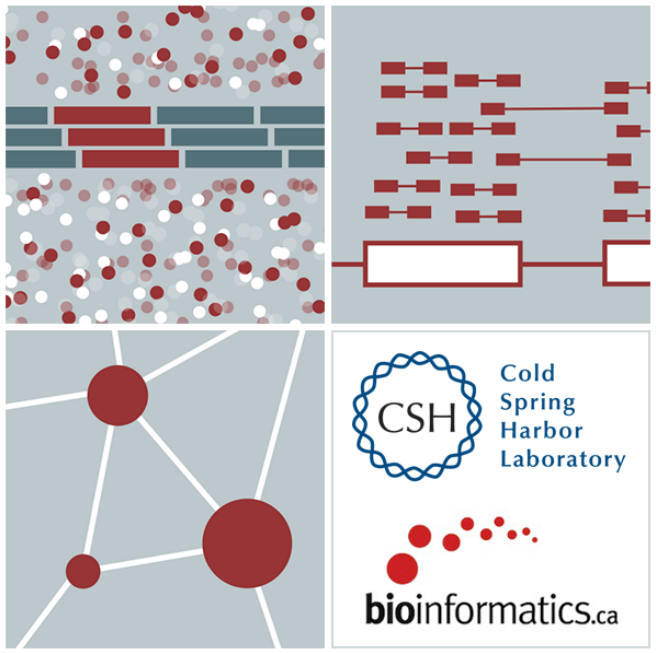
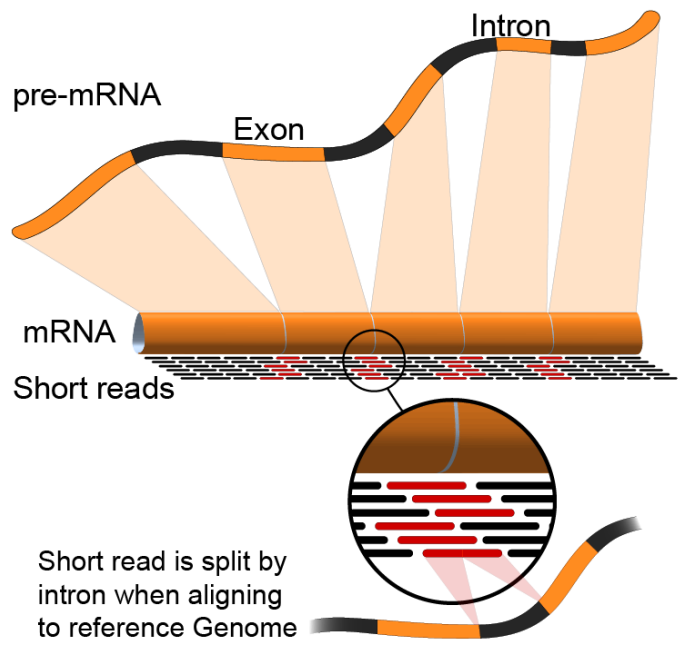
Introduction to single cell RNA sequencing and analysis

Kelsy Cotto, Malachi Griffith, Obi Griffith,
Evelyn Schmidt, Kartik Singhal, Zach Skidmore

CRI Bioinformatics Workshop. May 17-22, 2025





Washington University in St. Louis
School of Medicine

Acknowledgements

- Jon Preall, CSHL, SEQTEC
- Jennifer Foltz, Washington University, Genomics in Medicine
- Trevor Pugh, PMH, OICR, CBW
- Allegra Petti, MGH/Harvard
- Brian Haas, Broad Institute





Single-cell omics - A new paradigm in biology and medicine?

- Redefine “cell type”
  - Catalog known cell types in healthy and diseased tissues
  - Define/discover new cell types
- Redefine relationships between cells
- Interpret genetic variation at single-cell resolution
- And much more…

Discussion of bulk vs single cell RNA-seq

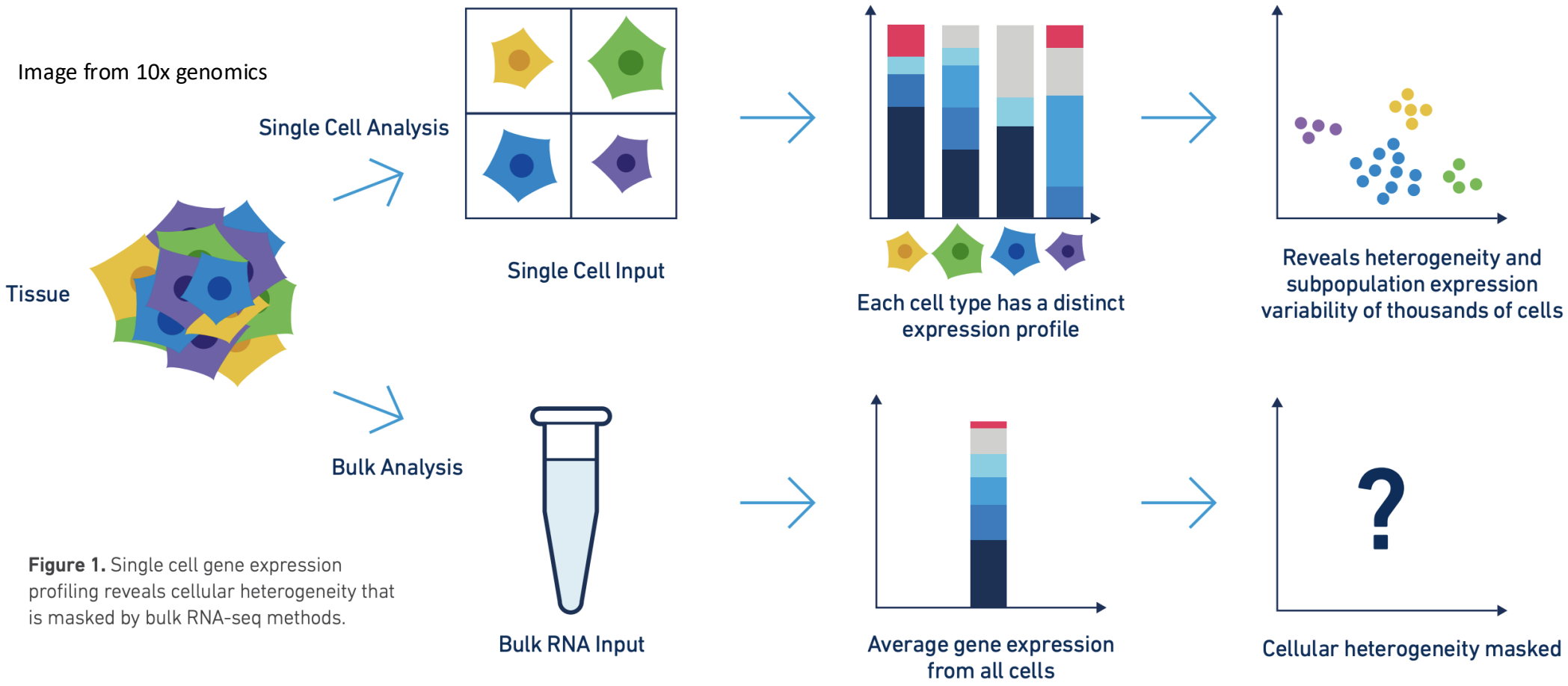
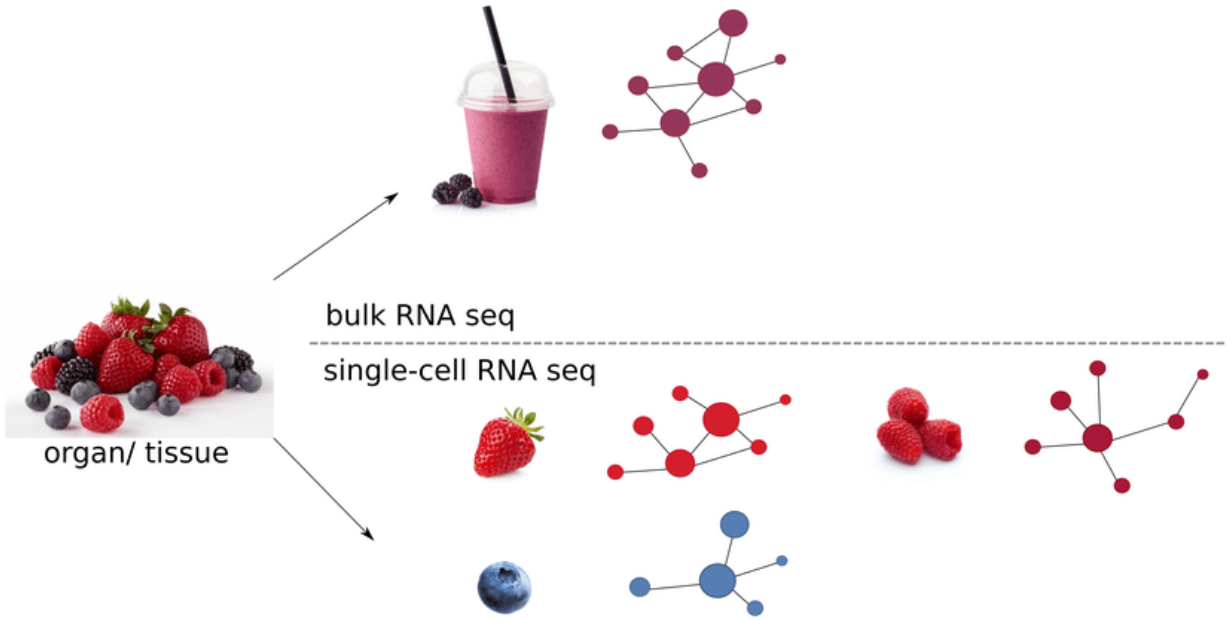


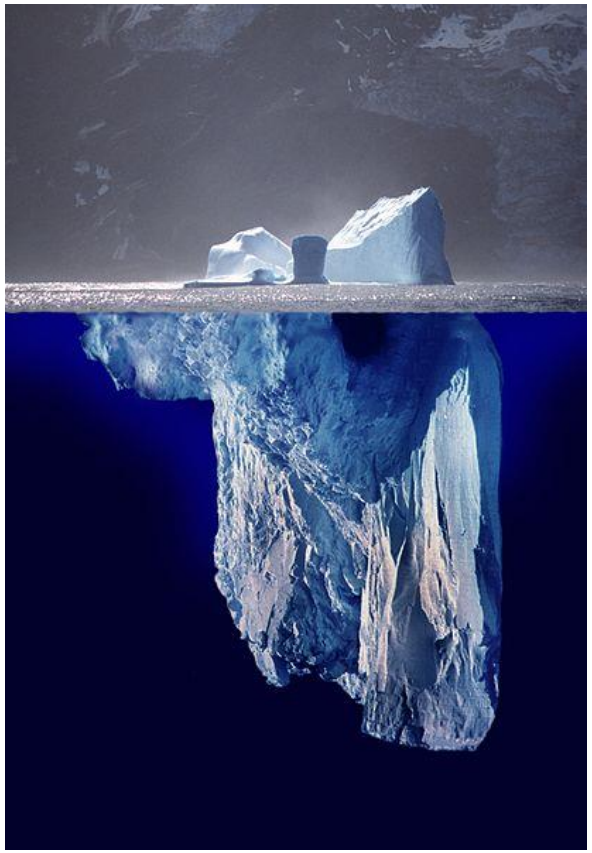
Figure 1. Single cell gene expression profiling reveals cellular heterogeneity that is masked by bulk RNA-seq methods.

What are key differences?

Factors to compare: Cost, complexity of library prep, complexity of analysis, qualitative and quantitative differences in richness of information obtained.

Fruit salad or tip of the iceberg?

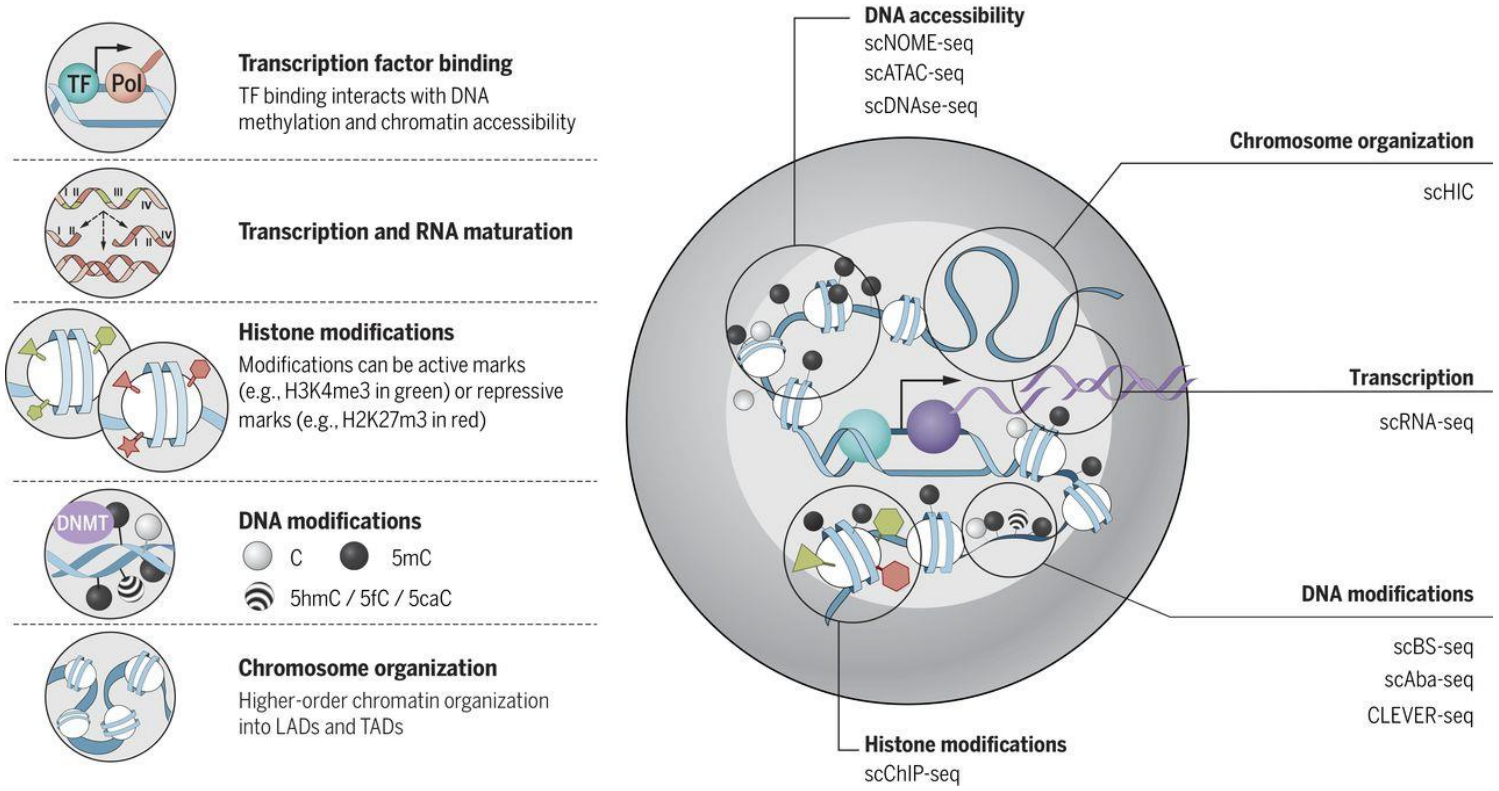




scRNA

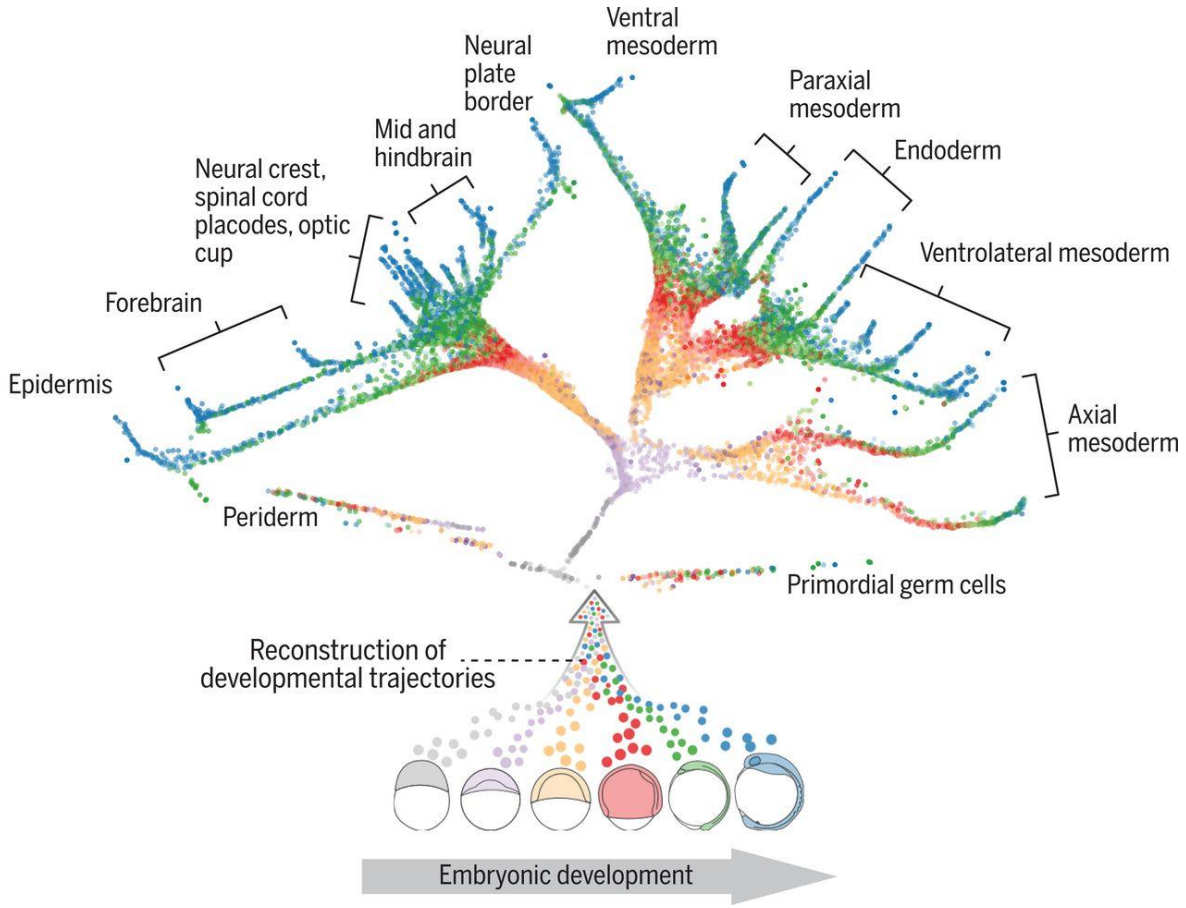
Multimodal single-cell measurements

- scRNA/snRNA
- CITE-seq
- TCR sequencing
- scATAC-seq
- scMethyl-seq
- TARGET-seq, G&T-seq (scDNA & scRNA)
- scCRISPRi/Perturb-seq
- Patch-seq
- Spatial transcriptomics



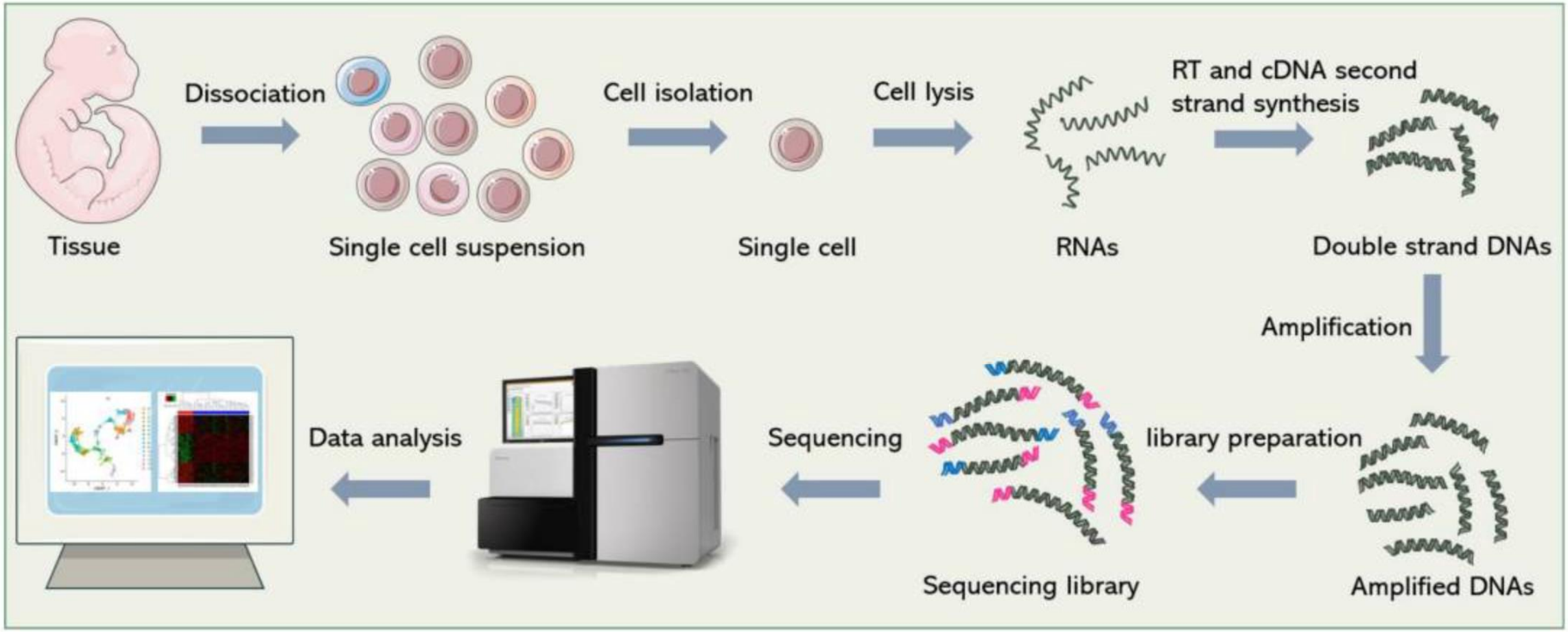
Kelsey et al. Science 2017

Cell type vs cell lineage vs cell state

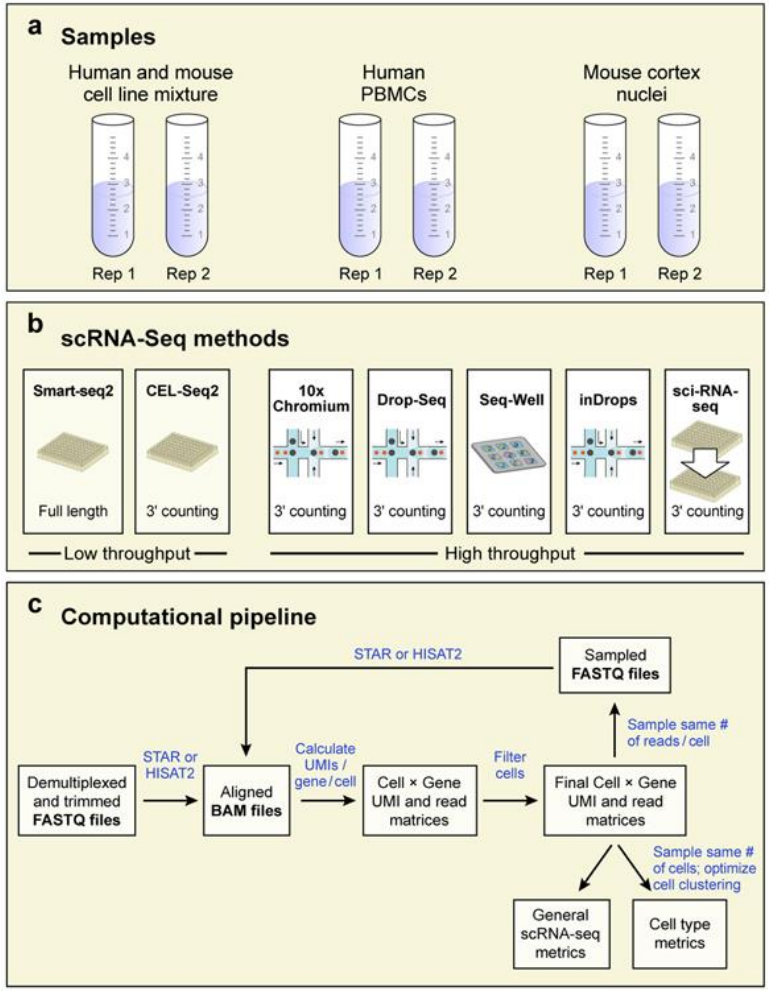


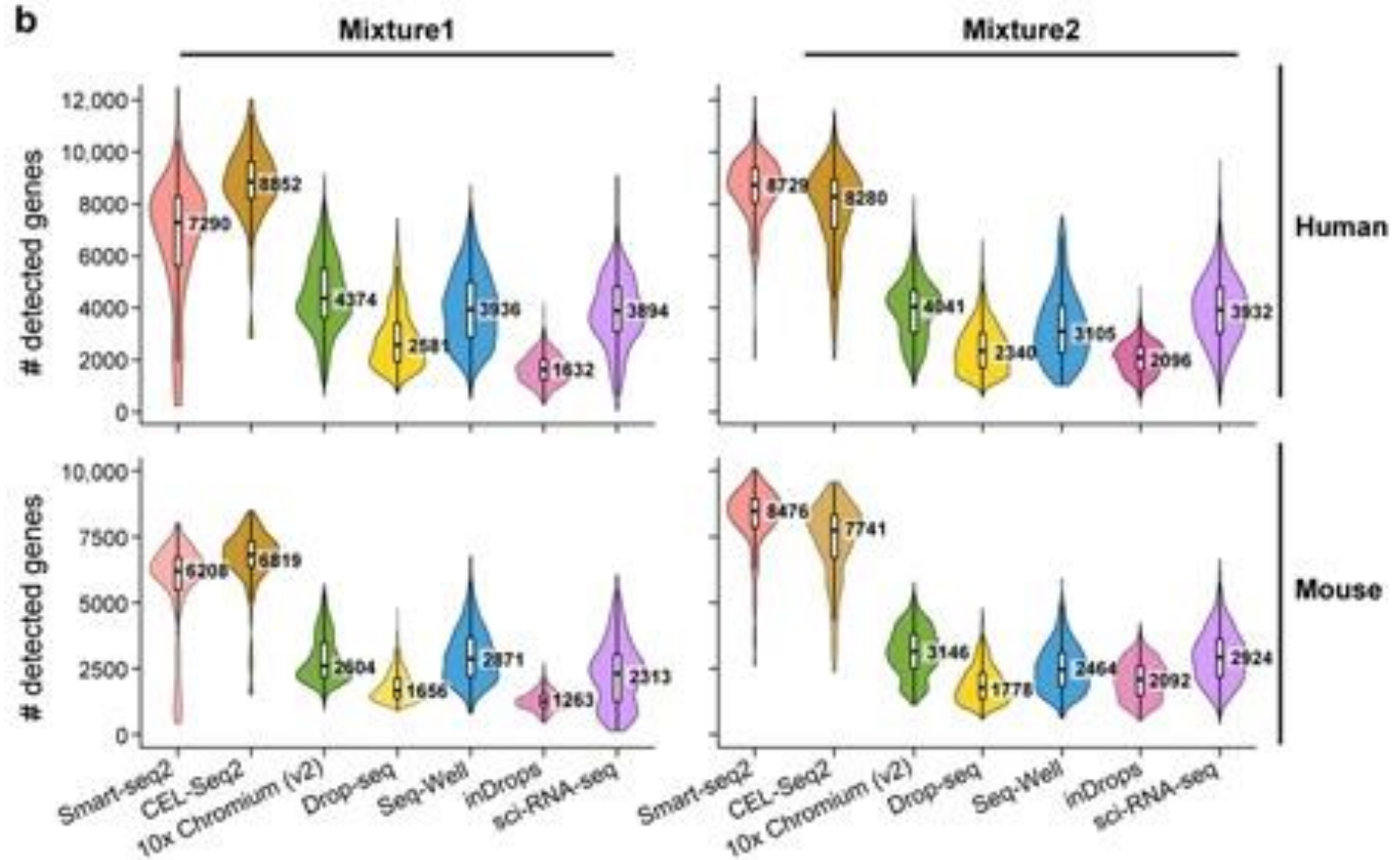
https://www.science.org/doi/10.1126/science.aar3131

Basic overview of single cell RNA analysis



Multiple platforms have emerged … Each with trade-offs…

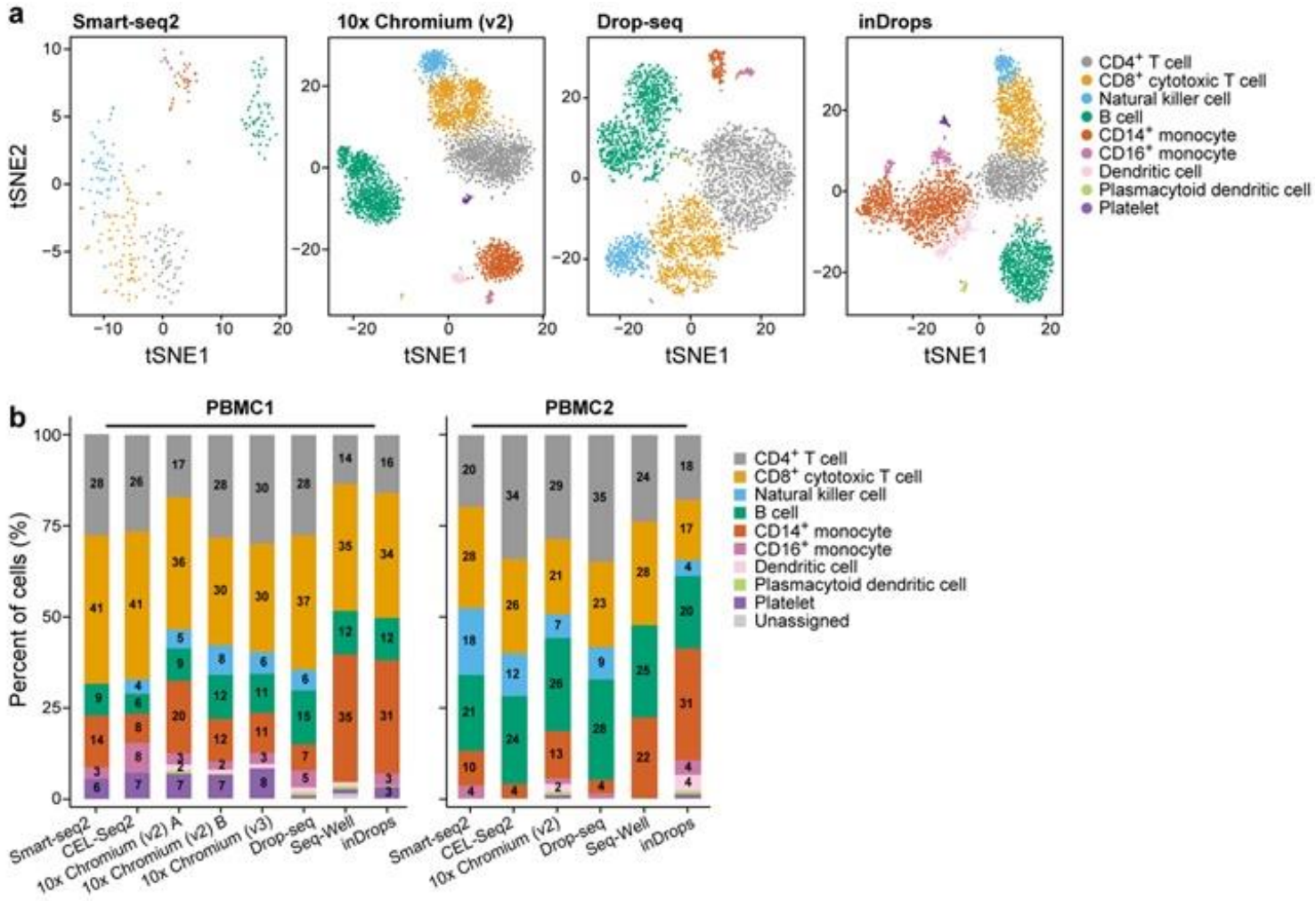




[PMC7289686]

Plate-based vs Droplet-based

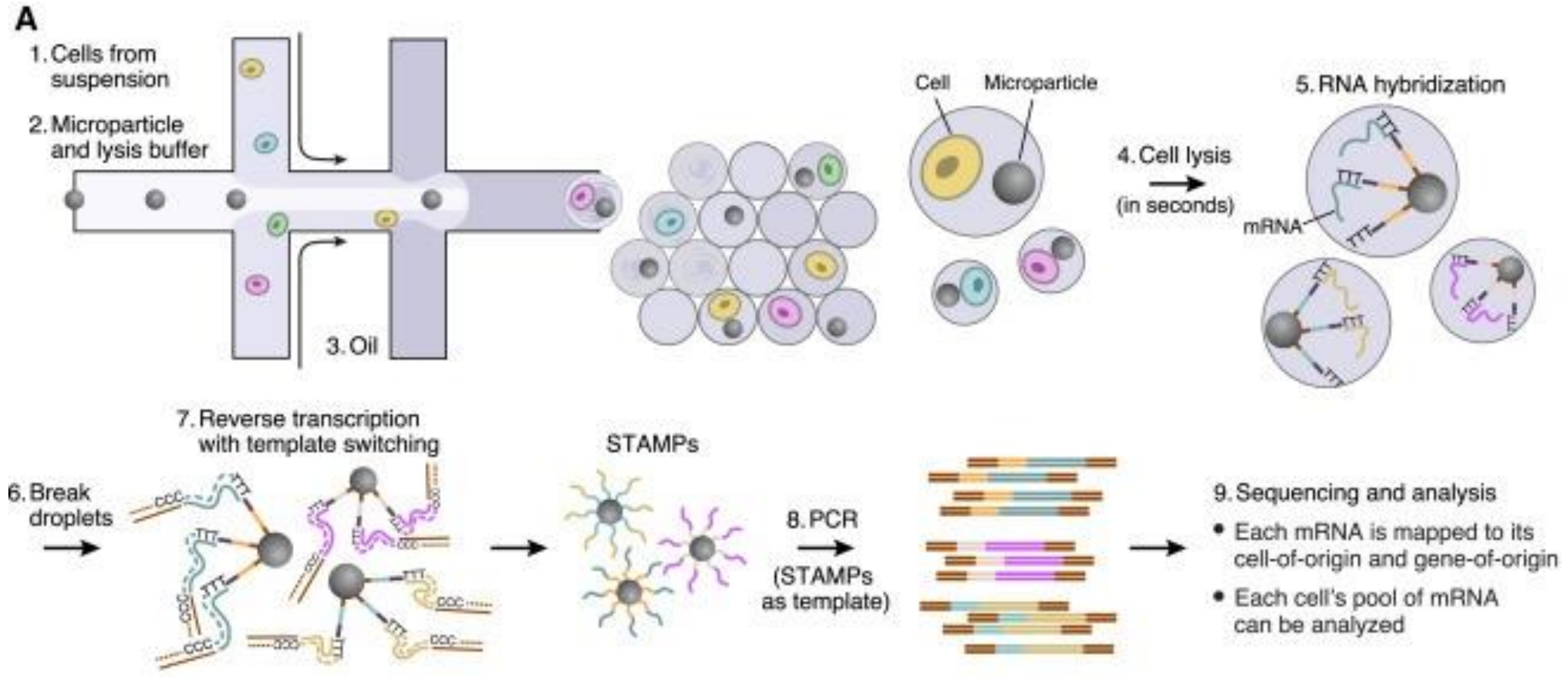
Multiple platforms have emerged … Each with trade-offs…



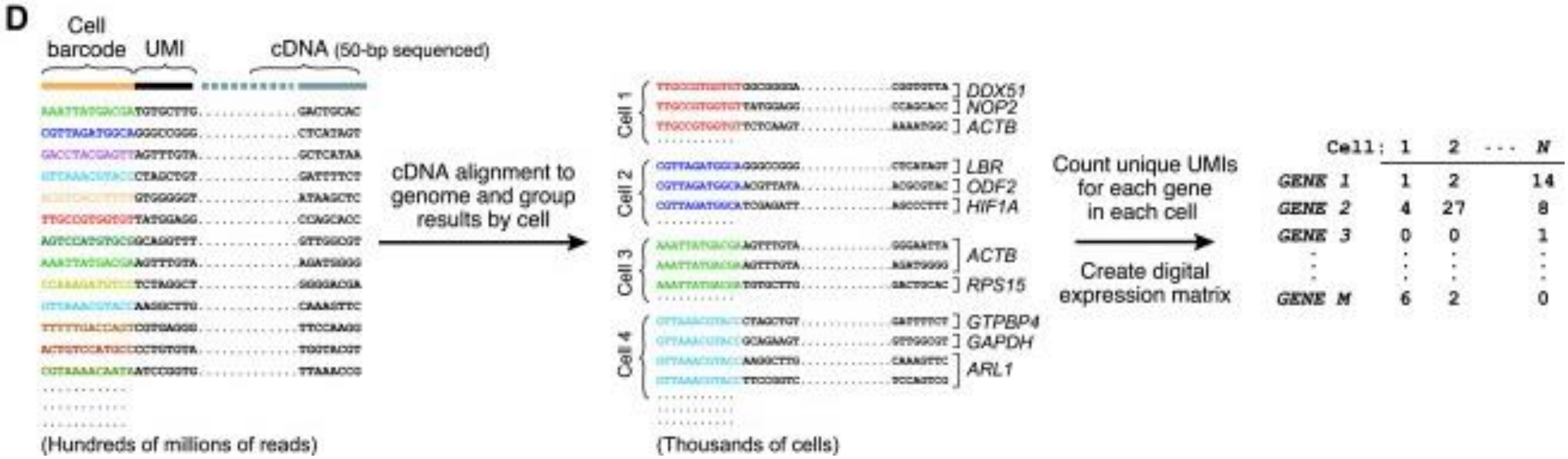
Additional considerations

- Size of cells
- Expected % of cell population of interest
- Depth of detection needed for experimental question
- Accessibility

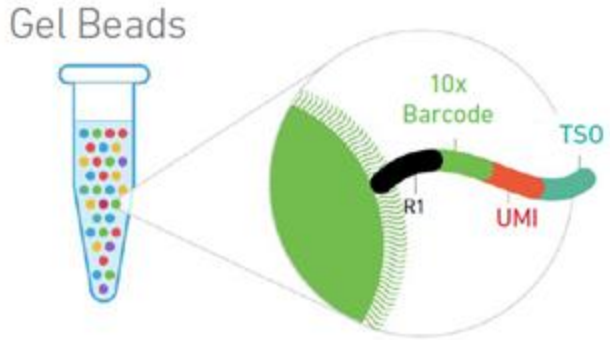
Intro to 10X approach

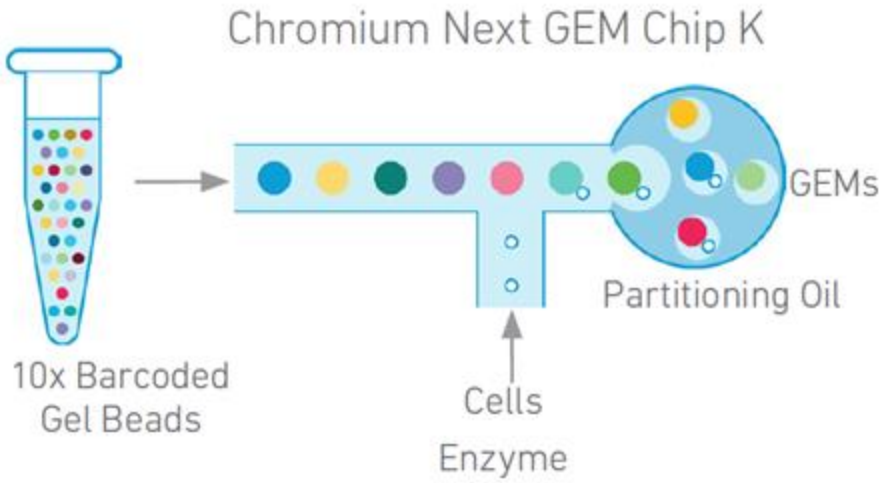


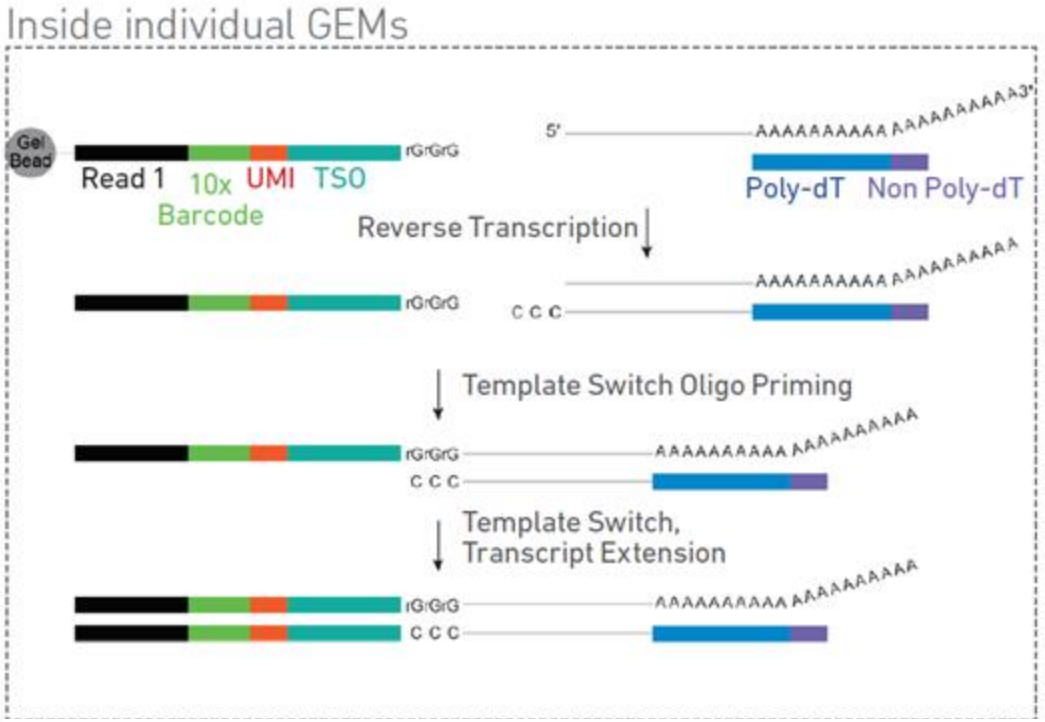
Intro to 10X approach

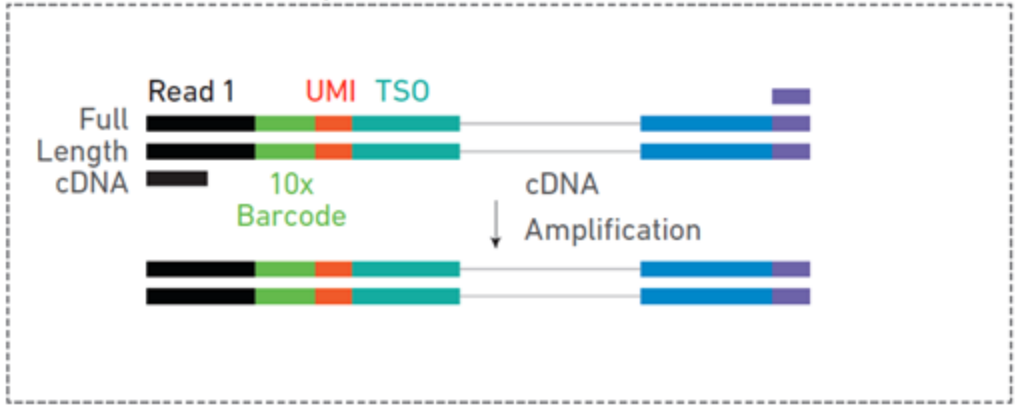


10X Approach (5’)

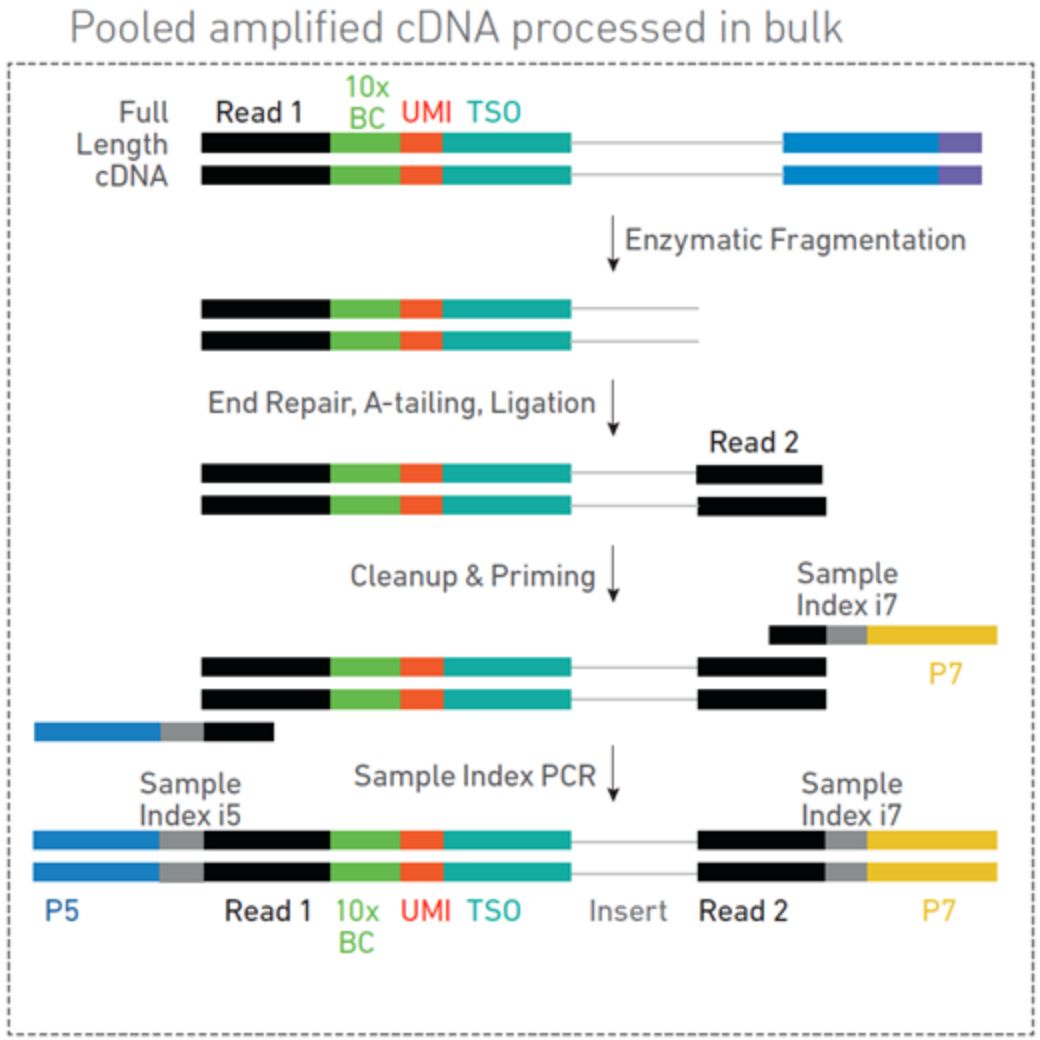


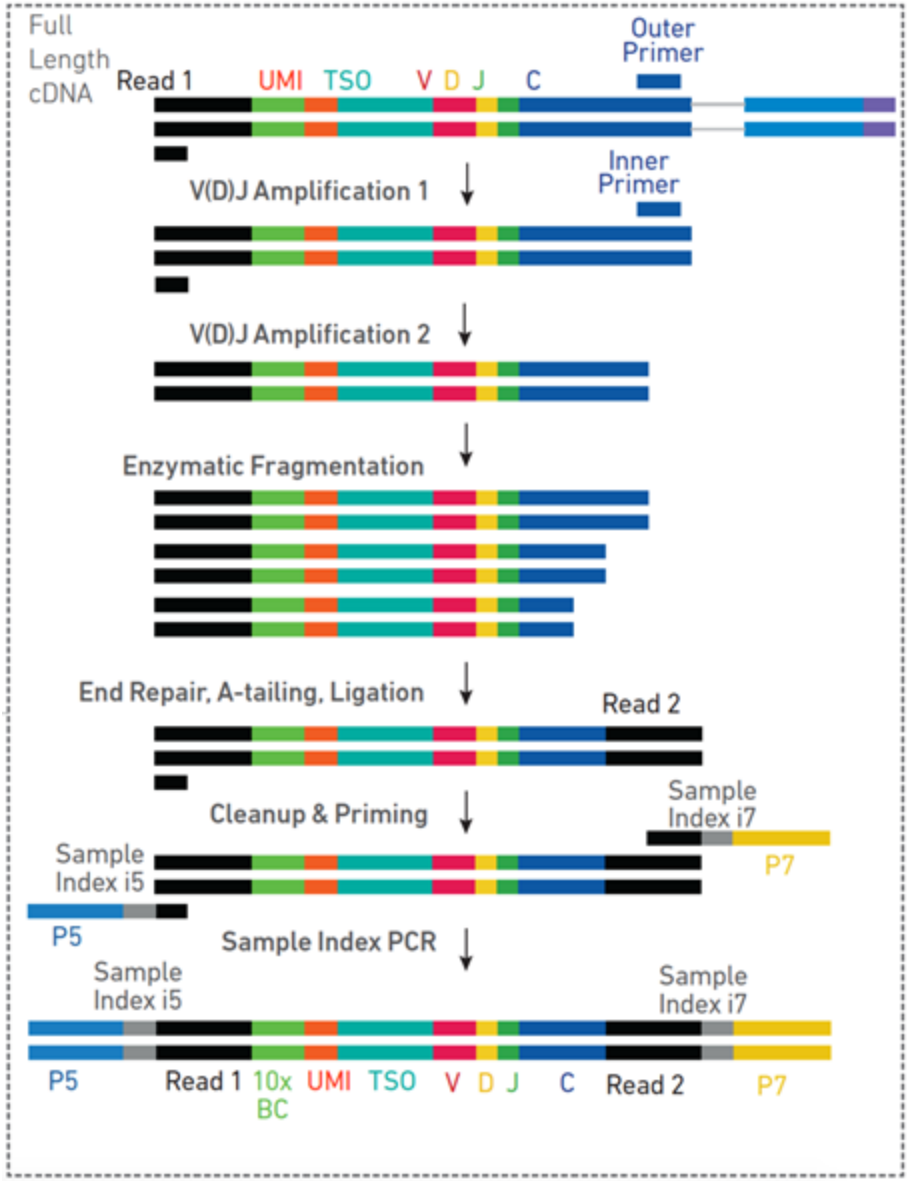






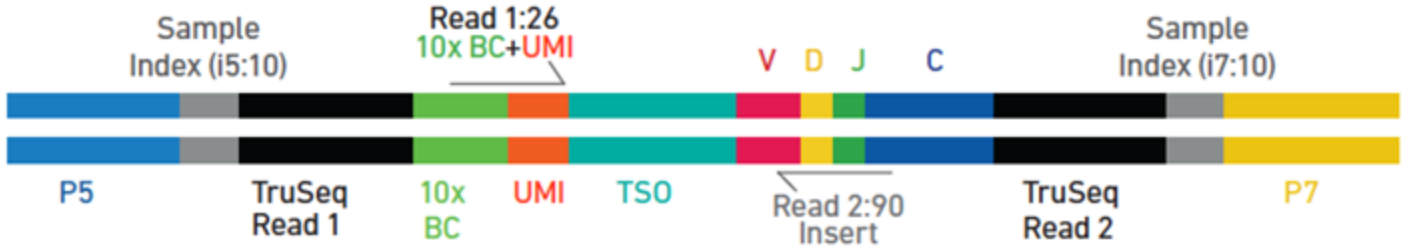
10X Approach (5’ v2)





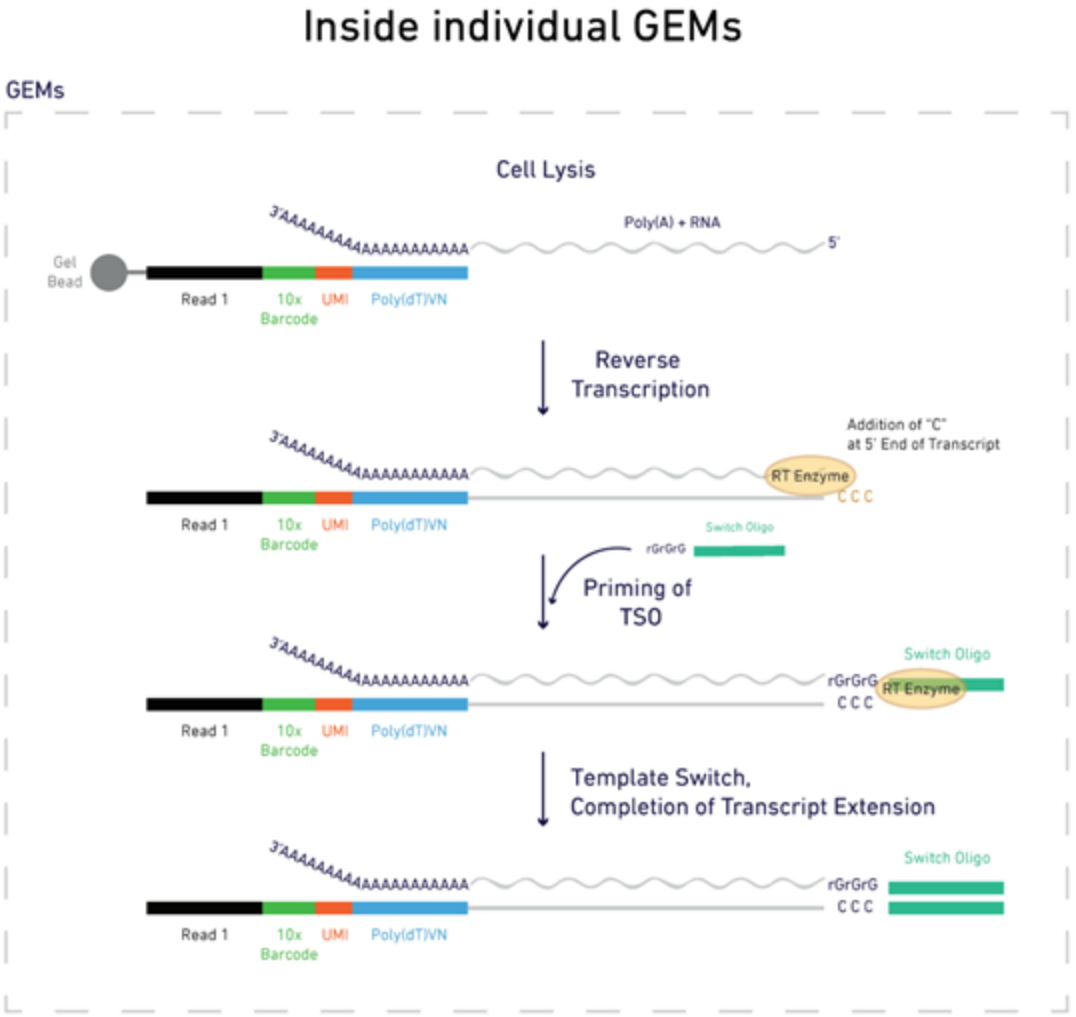
10X Single Cell 5’ GEX and V(D)J dual index libraries

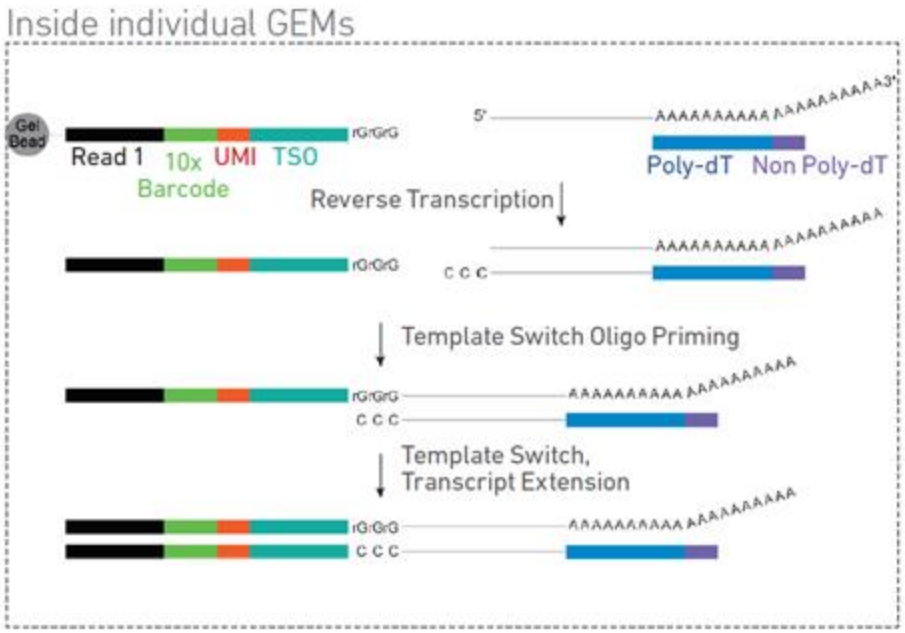
Chromium Single Cell V(D)J Dual Index Library



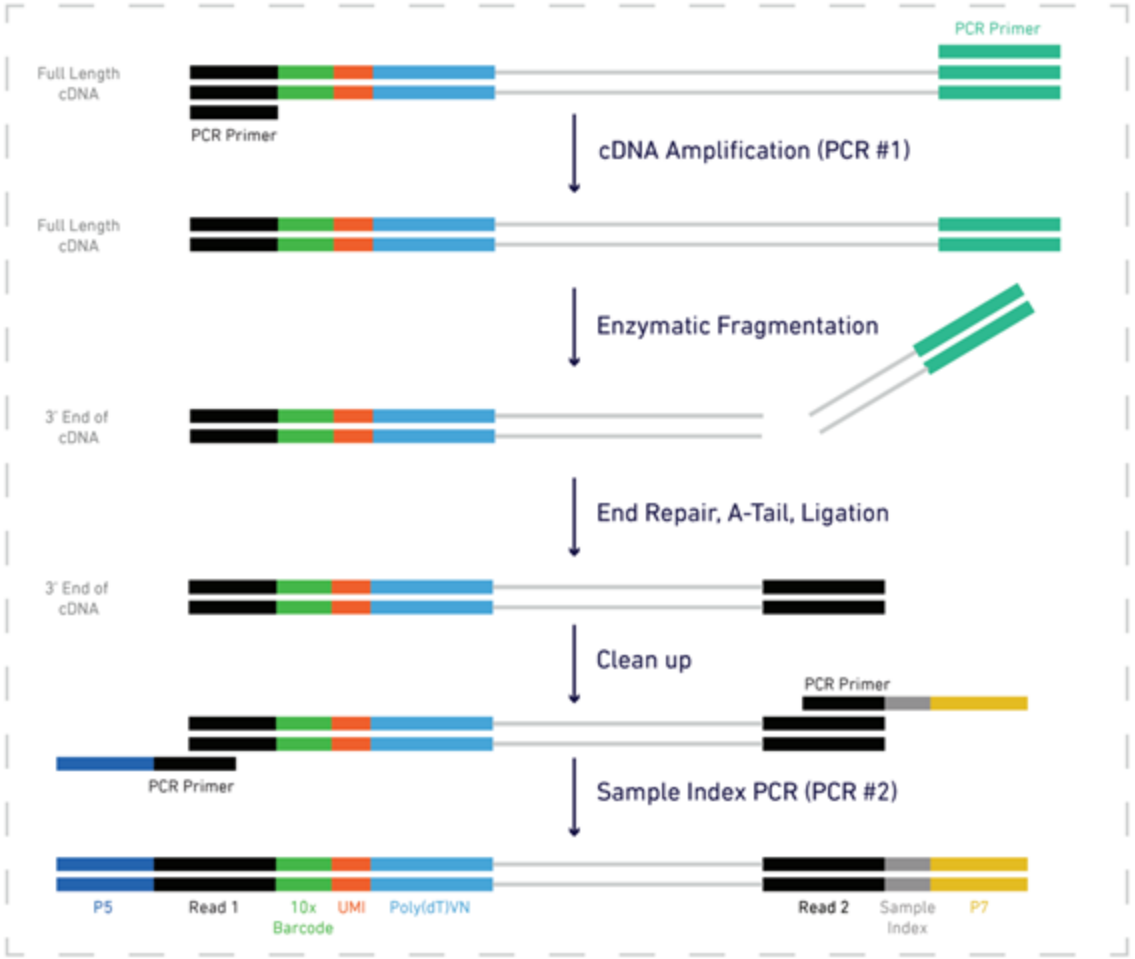
Chromium Single Cell 5' Gene Expression Dual Index Library

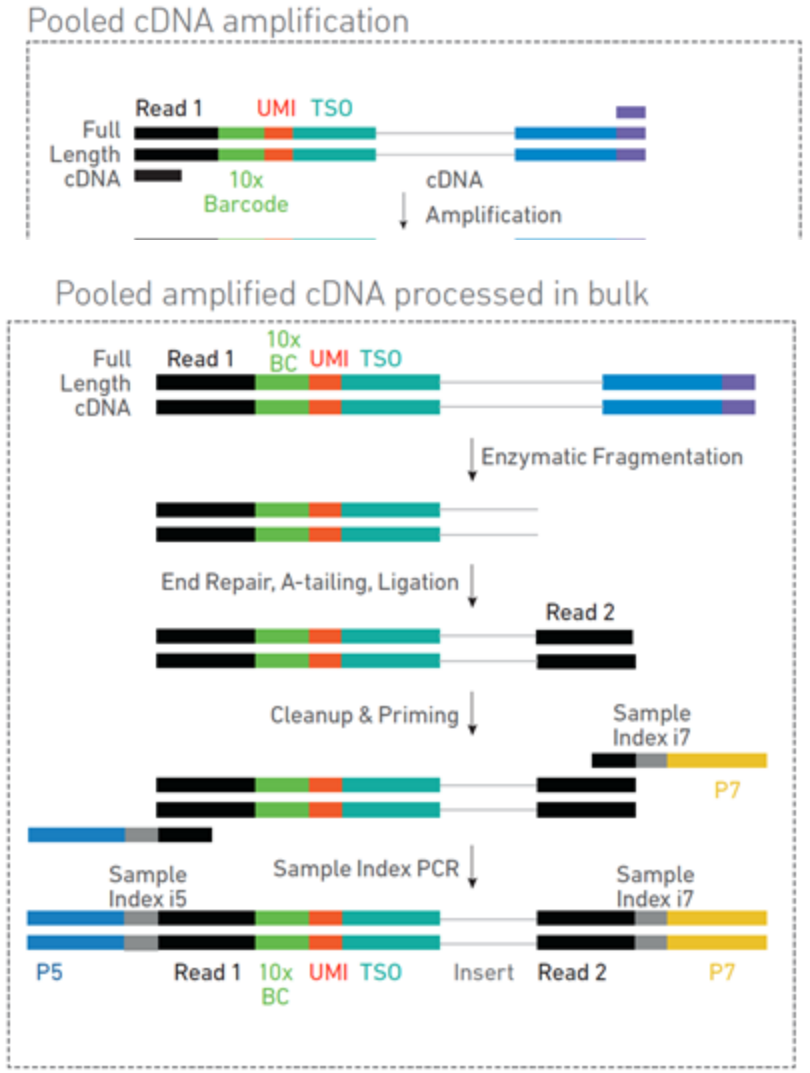
3’ vs 5’ approach - polyT and TSO switch places



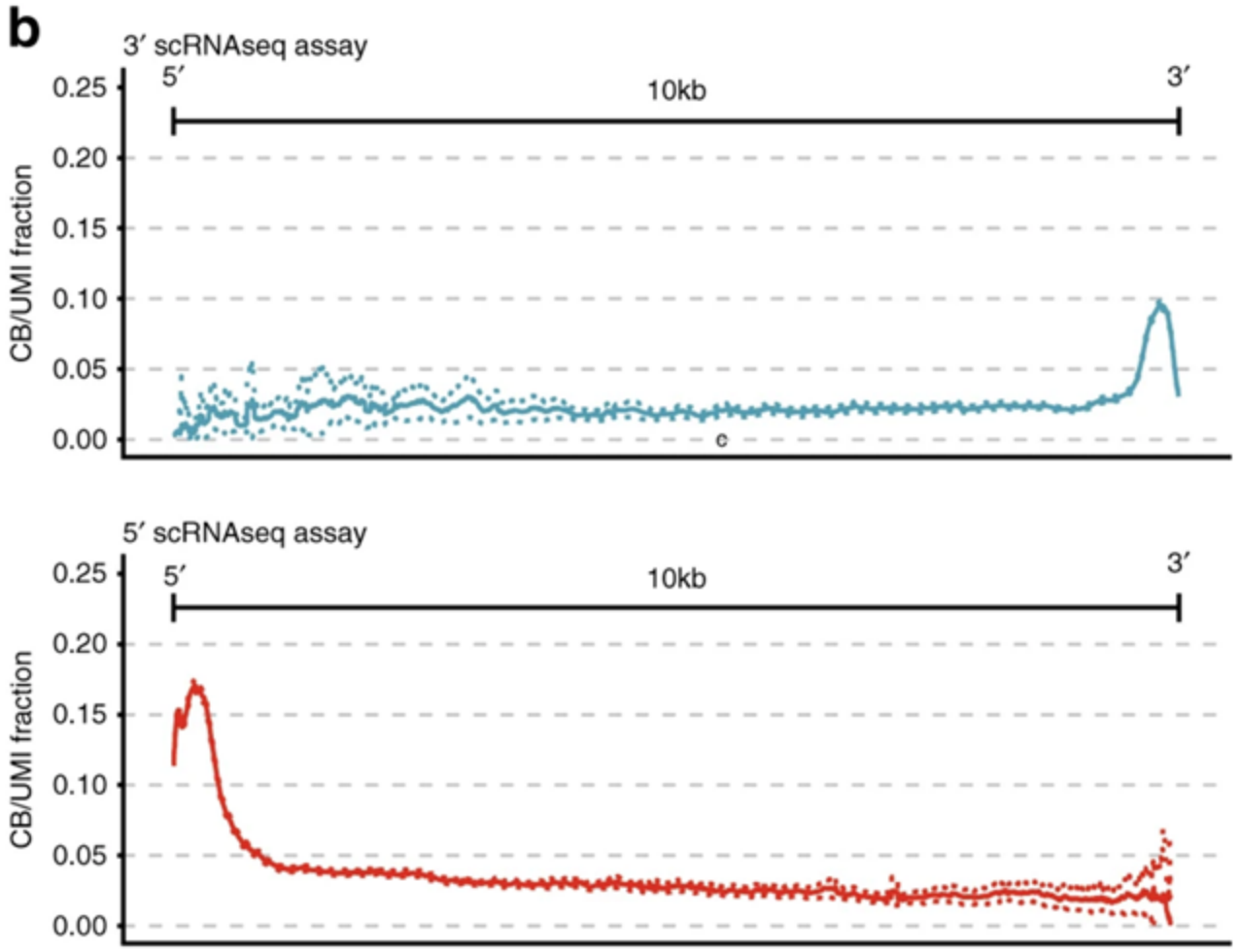


3’ vs 5’ approach - Fragmentation and PCR steps result in different ends of transcripts being kept





3’ vs 5’ approach results in different end bias



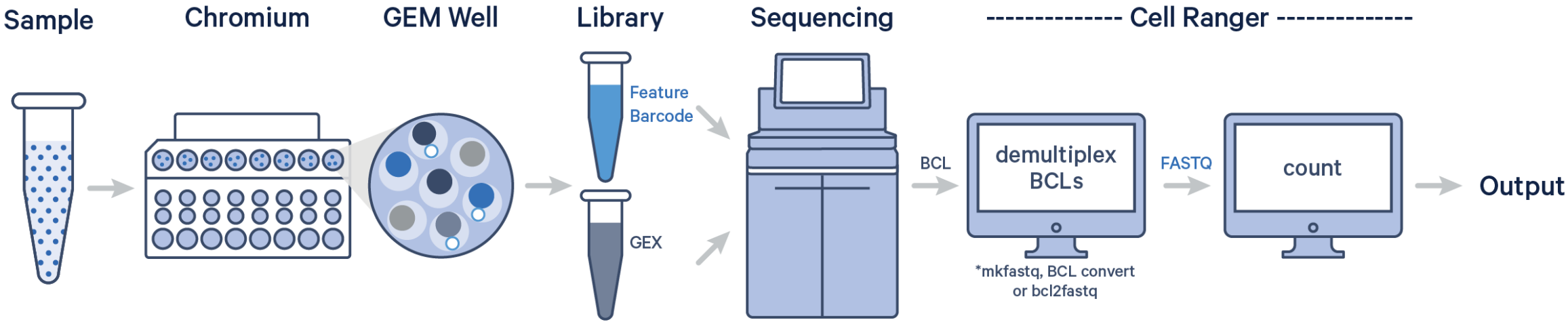
PMC6694122

Fig1b. Petti et al. 2019. Nat Comm

3’ vs 5’ approach

- Both require polyA transcripts
- Both involve end bias (choose 3’ or 5’)
  - 5’ coverage more evenly distributed?
- Only 5’ approach is compatible with V(D)J enrichment
  - Because business end of V(D)J transcript is at 5’ followed by large C region at 3’ end
  - In order to amplify TCR from 3’ 10X library you would need very large fragments and would have to design primers off the variable V genes instead of constant C gene

Overview of 10x genomics pipeline

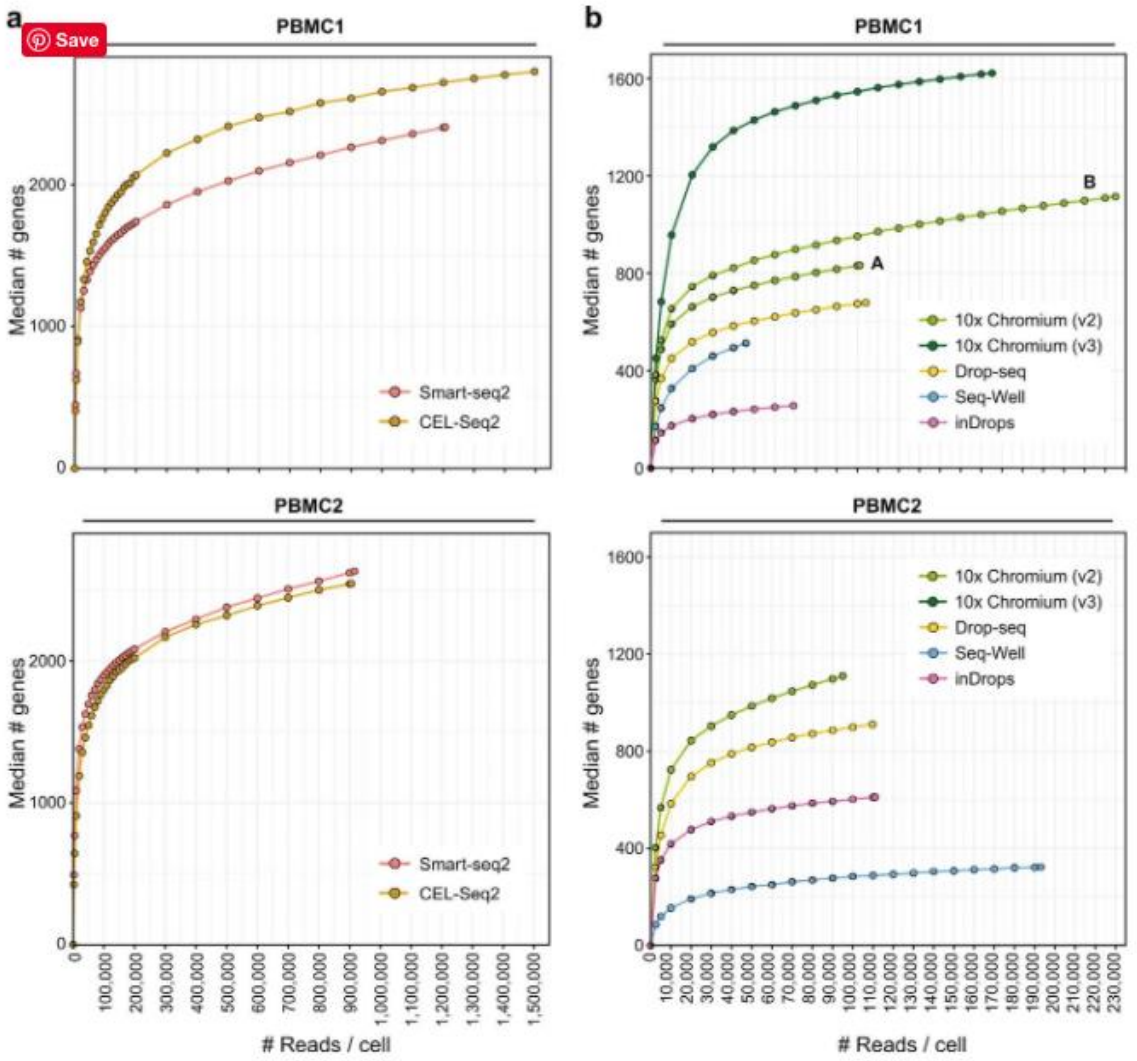


How deeply do you need to sequence?

General Rule:
Achieve 90% saturation

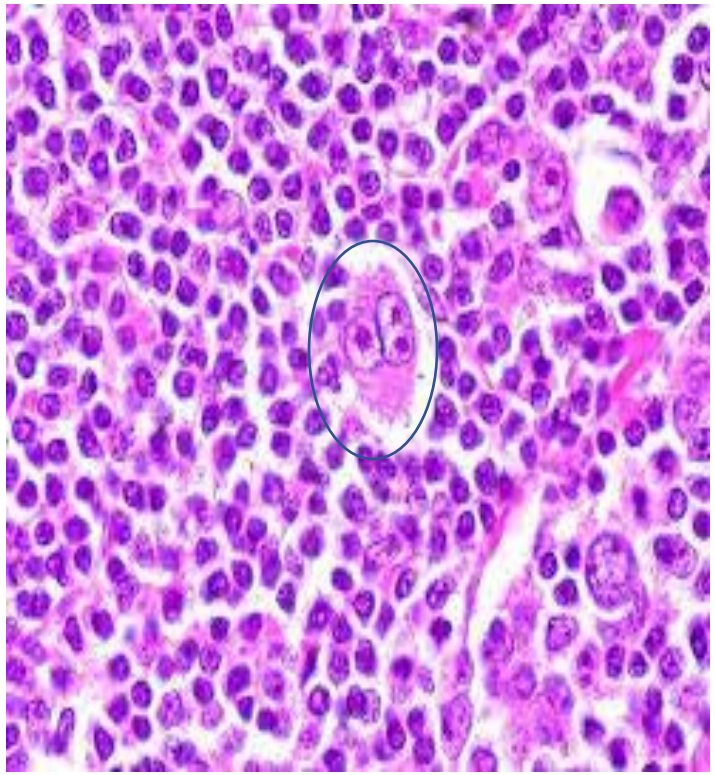
Official Recommendations (reads/cell):

3’ - V3: 20K
3’ - V2: 50K
5’ - 20K
5’ with variant discovery - 200K
5’ V(D)J - 5K
Fixed RNA: 10K



Example application: Increased capacity and new technologies are allowing study of previously inaccessible tumor types

- Hodgkins Lymphoma - has remained a challenge because HRS cells account for ~1-5% of cells in the tumor tissue.



- Isolate HRS cells and then apply genomic techniques
  - Flow sorting – very challenging cell type
  - Laser capture microdissection – extremely low inputs for sequencing
- Handful of cell lines have been profiled
- Almost no genome-wide sequencing data exists

Todd Fehniger

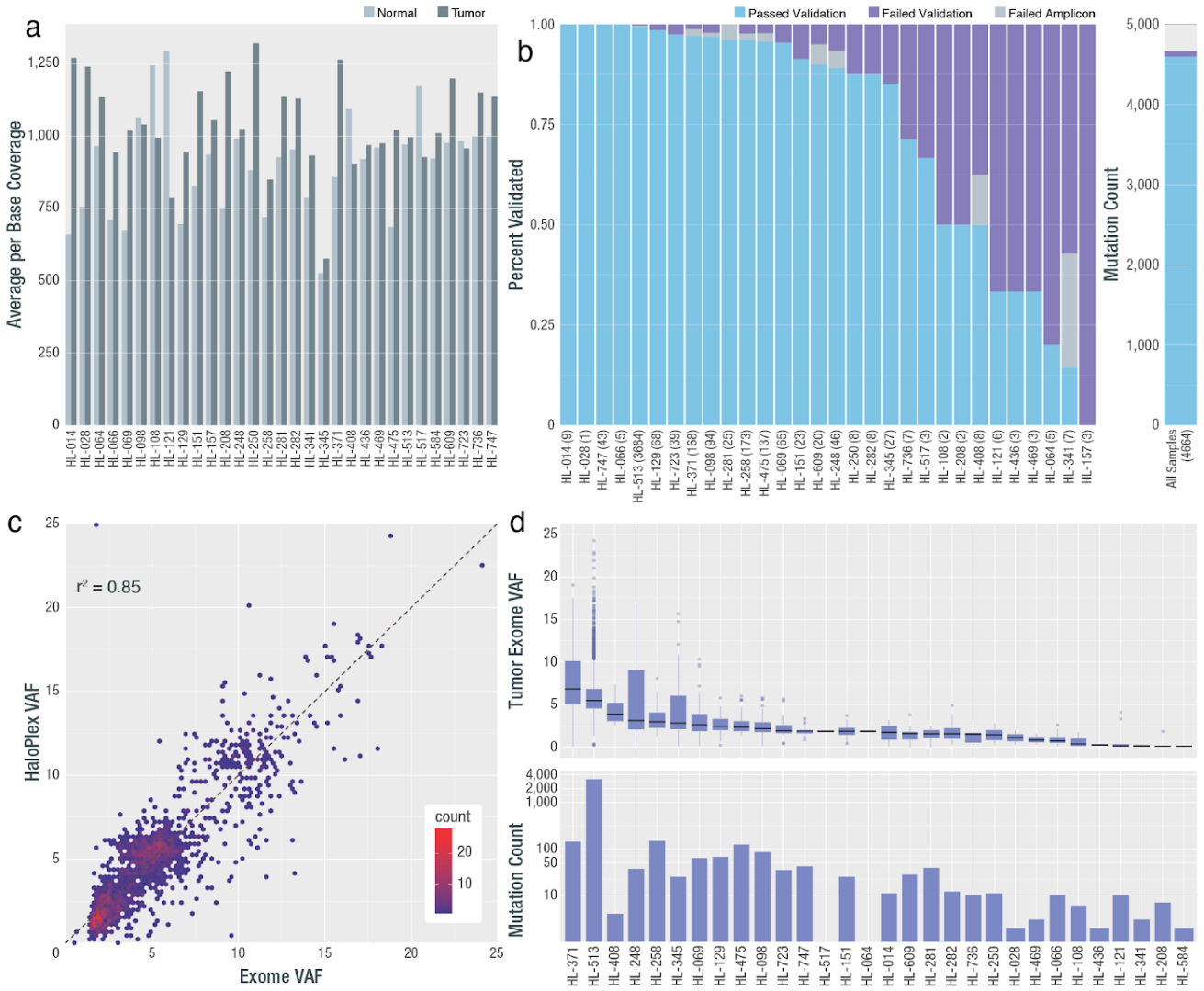
Felicia Gomez

Brute force strategy - Ultra Deep Exome sequencing of 31 Hodgkins tumor/normal pairs

- IDT Exome capture reagent
- Three KAPA libraries were constructed/sample
- Libraries were sequenced across eight lanes of an Illumina HiSeq
- Somatic Variant Calling
  - SNV were called using 5 variant callers
  - Indels were called using 4 variant callers
- Target depth ~1000x

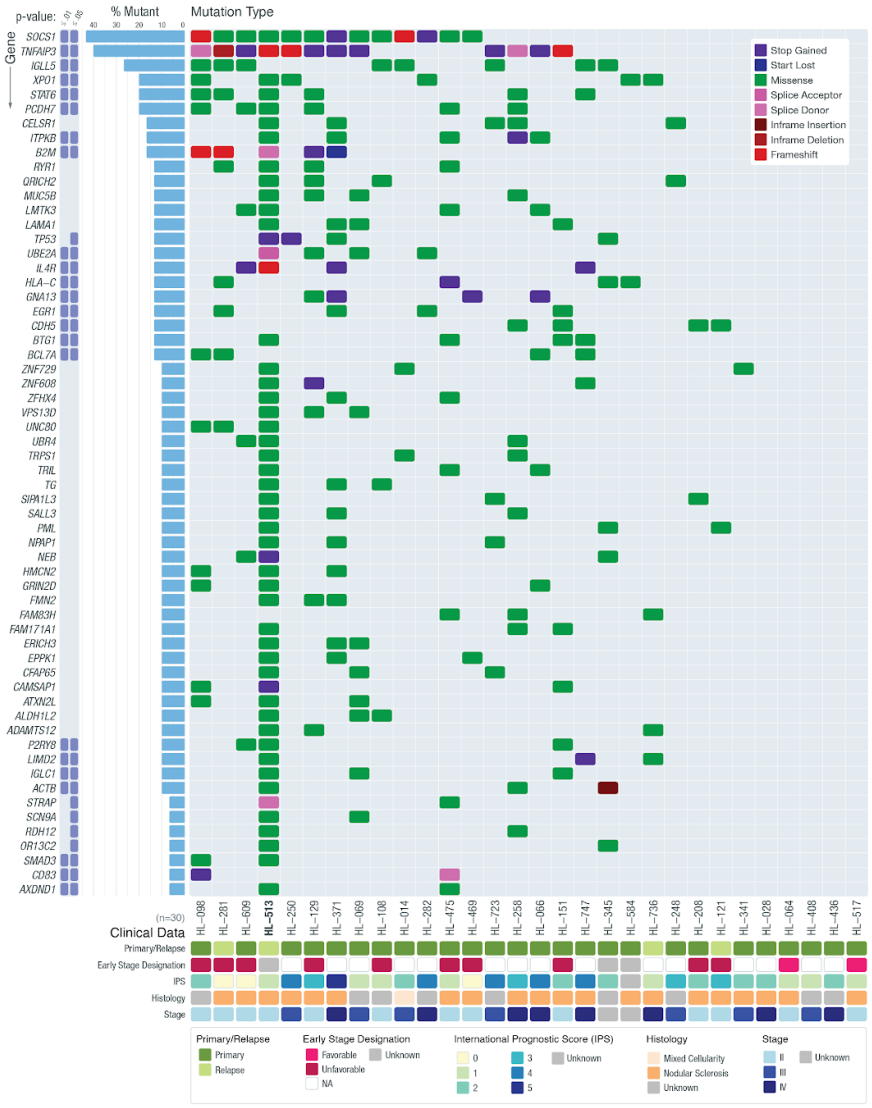
- Validation with targeted Haloplex technology

High coverage and concordance achieved

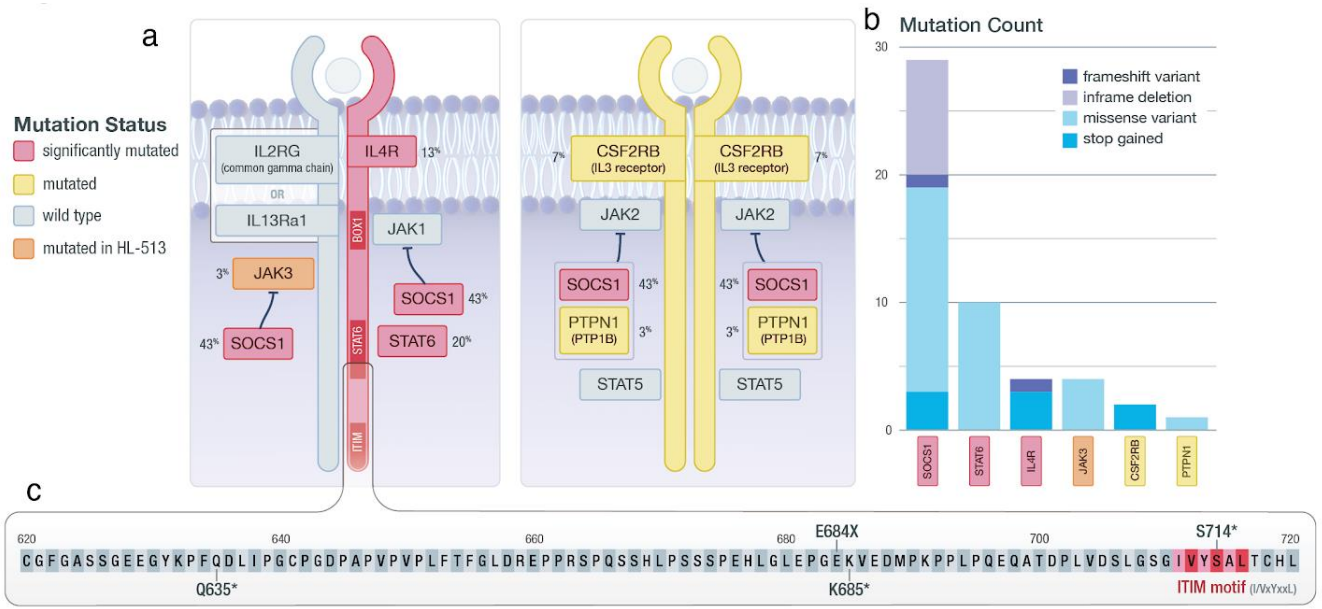


- ~1000X coverage achieved
- High VAF concordance between WES and haloplex validation
- Overall variant validation rate >90%
- Mean mutation burden = 33 protein-coding variants
  - One hypermutator 3,160 variants

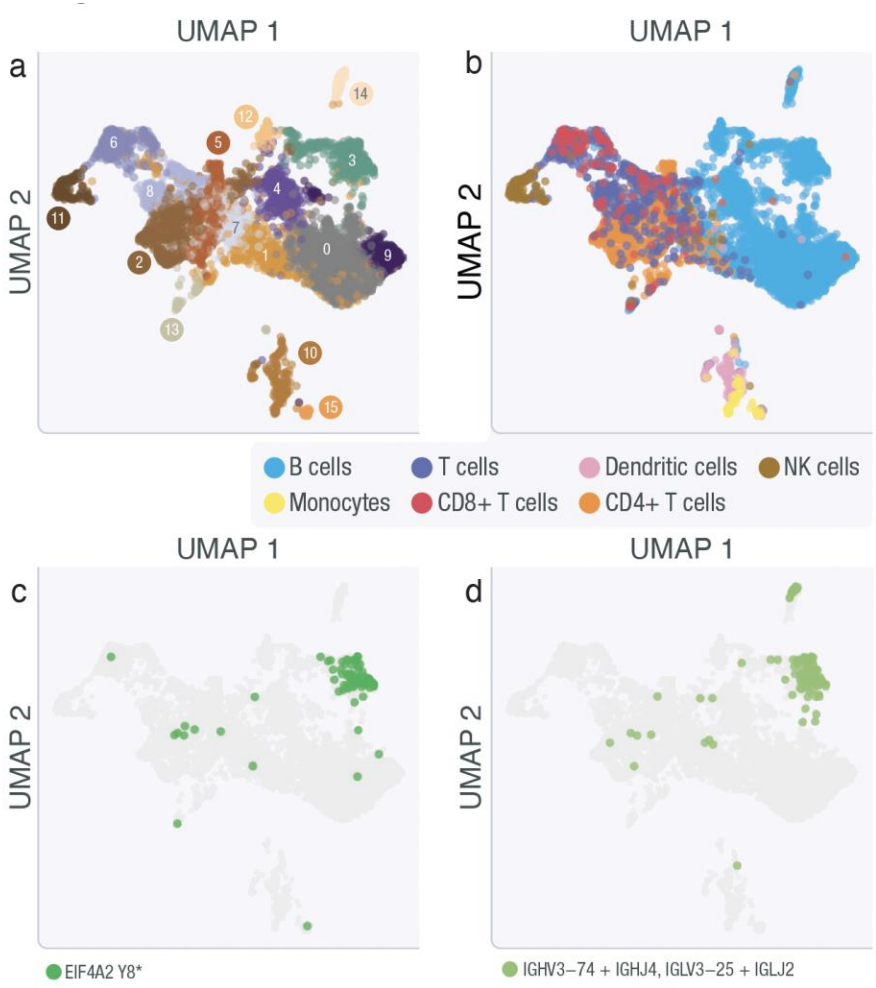
Represents one of the few comprehensive surveys of HL - mutation landscape reveals known and novel genes/pathways/hotspots

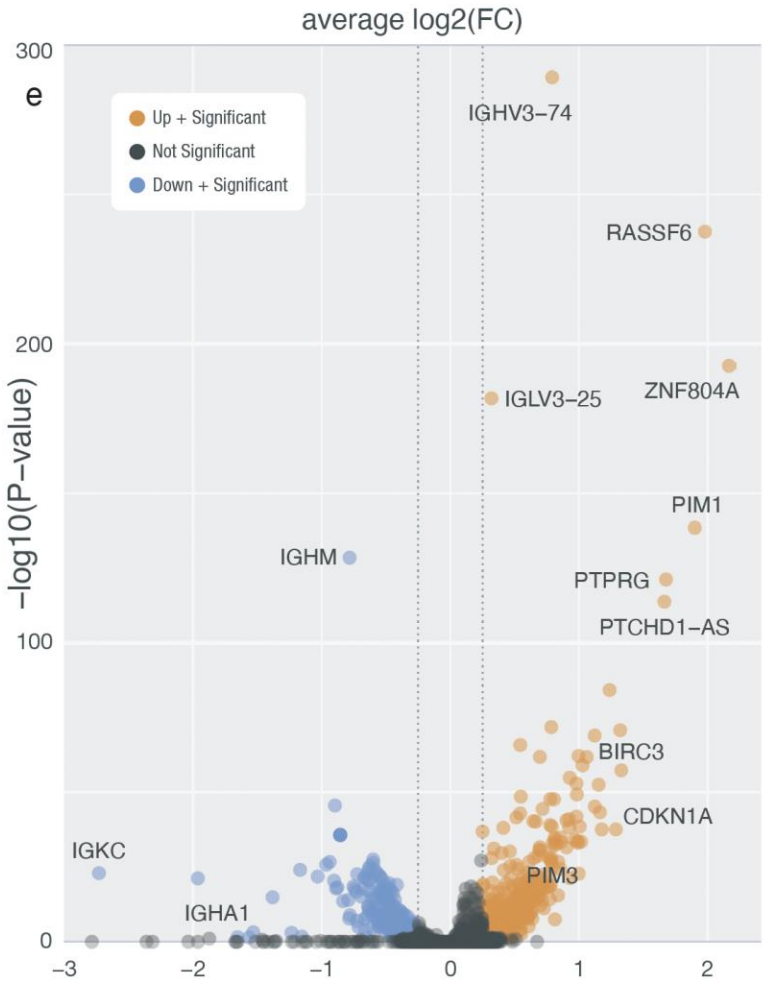


- ~50 genes have mutations in 3 or more samples
- *JAK/STAT - SOCS1 & STAT6*
- *NFKB - TNFAIP3 & XPO1*
- *SWI/SNF - BCL7A & SMAD3*
- *PCDH7, IGLL5, CDH5, …*



Single nucleus sequencing identifies mutant-positive HRS cells





CANCER RESEARCH COMMUNICATIONS

AACR American Association for Cancer Research

RESEARCH ARTICLE https://doi.org/10.1158/2767-9764.CRC-23-0140 OPEN ACCESS

Ultra-Deep Sequencing Reveals the Mutational Landscape of Classical Hodgkin Lymphoma

Felicia Gomez[1,2,3], Bryan Fisk[1,2], Joshua F. McMichael[2], Matthew Mosior[1,2], Jennifer A. Foltz[1], Zachary L. Skidmore[1,2], Eric J. Duncavage[4], Christopher A. Miller[1,2], Haley Abel[1,2], Yi-Shan Li[4], David A. Russler-Germain[1], Kilannin Krysiak[1,2,3,4], Marcus P. Watkins[1], Cody A. Ramirez[1,2], Alina Schmidt[1,2], Fernanda Martins Rodrigues[1,2], Lee Trani[2], Ajay Khanna[1], Julia A. Wagner[1], Robert S. Fulton[2], Catrina C. Fronick[2], Michelle D. O'Laughlin[2], Timothy Schappe[1], Amanda F. Cashen[1], Neha Mehta-Shah[1], Brad S. Kahl[1], Jason Walker[2], Nancy L. Bartlett[1], Malachi Griffith[1,2,3,5], Todd A. Fehniger[1,3], and Obi L. Griffith[1,2,3,5]

Introduction to scRNA dataset for hands-on exercises

https://rnabio.org/module-08-scrna/0008/01/02/scRNA_Data/

Includes brief demo of CellRanger QC Reports and Loupe Browser