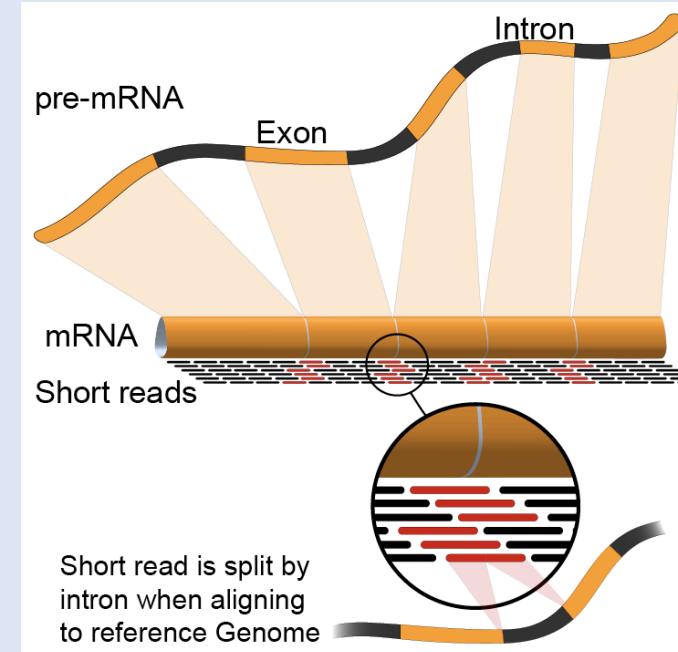
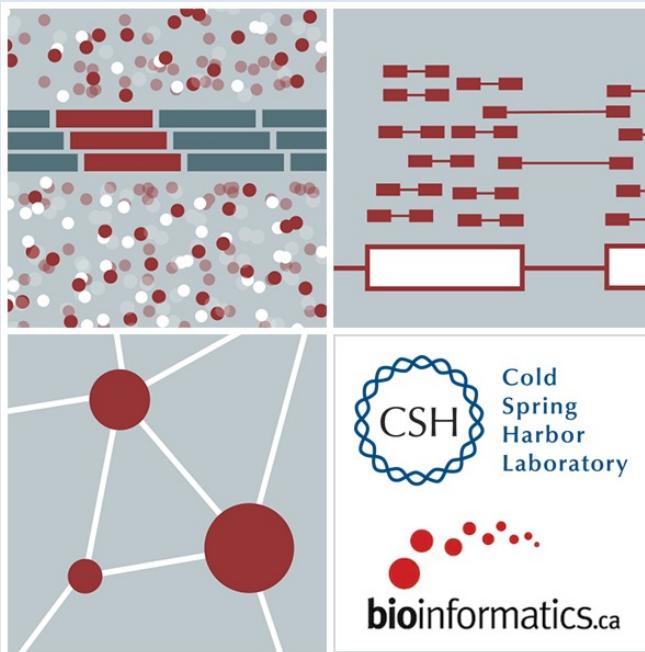




Cancer
Research
Institute™

Introduction to single cell RNA sequencing and analysis

Kelsy Cotto, Malachi Griffith, Obi Griffith, Charles Hayfron-Benjamin, Evelyn Schmidt, Kartik Singhal, Zach Skidmore
CRI Bioinformatics Workshop. Apr 27-May 2, 2024



 Washington University in St. Louis
SCHOOL OF MEDICINE

Acknowledgements

- Jon Preall, CSHL, SEQTEC
- Jennifer Foltz, WashU, Genomics in Medicine
- Trevor Pugh, PMH, 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

Discussion of bulk vs single cell RNA-seq

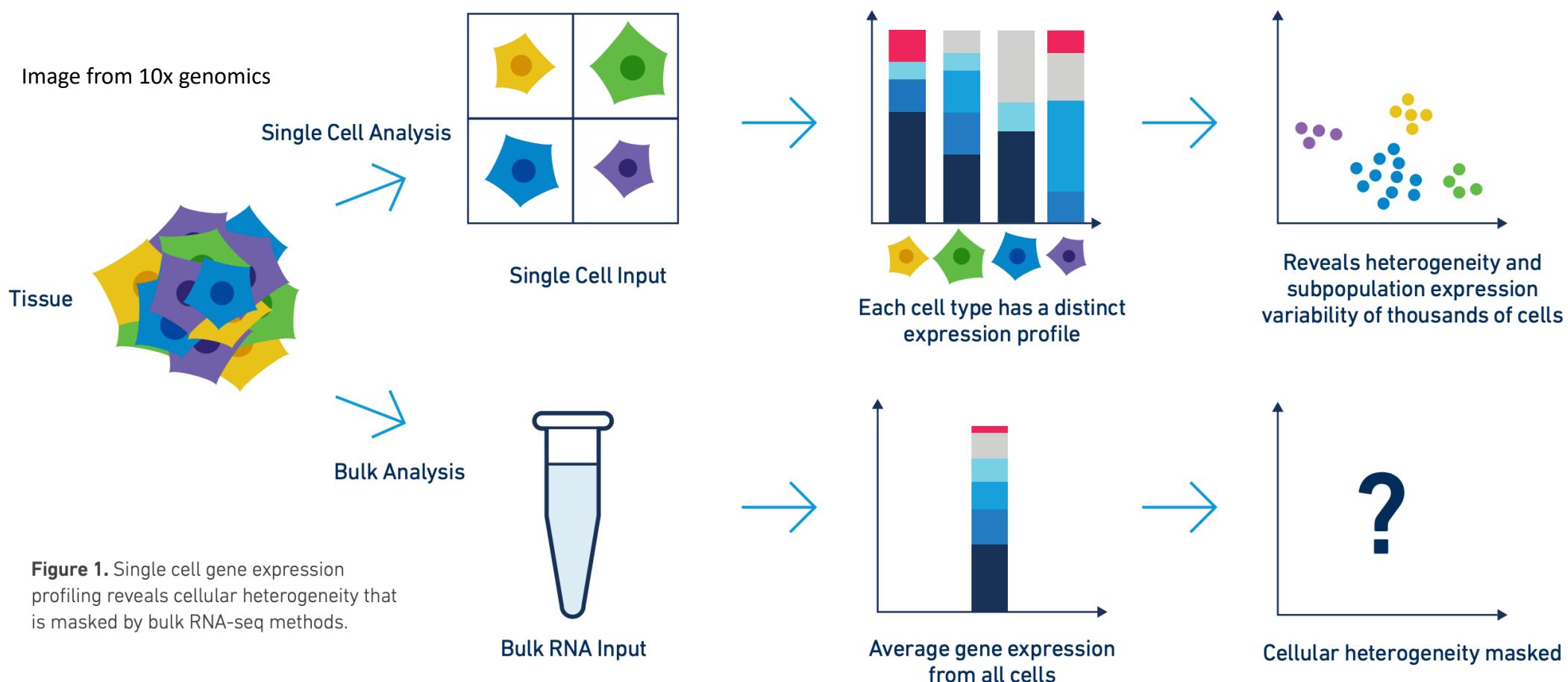
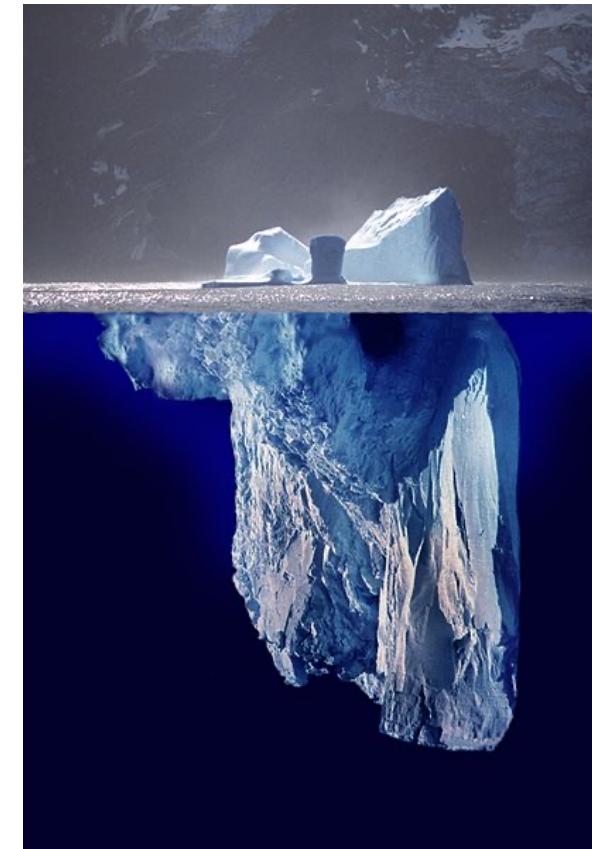
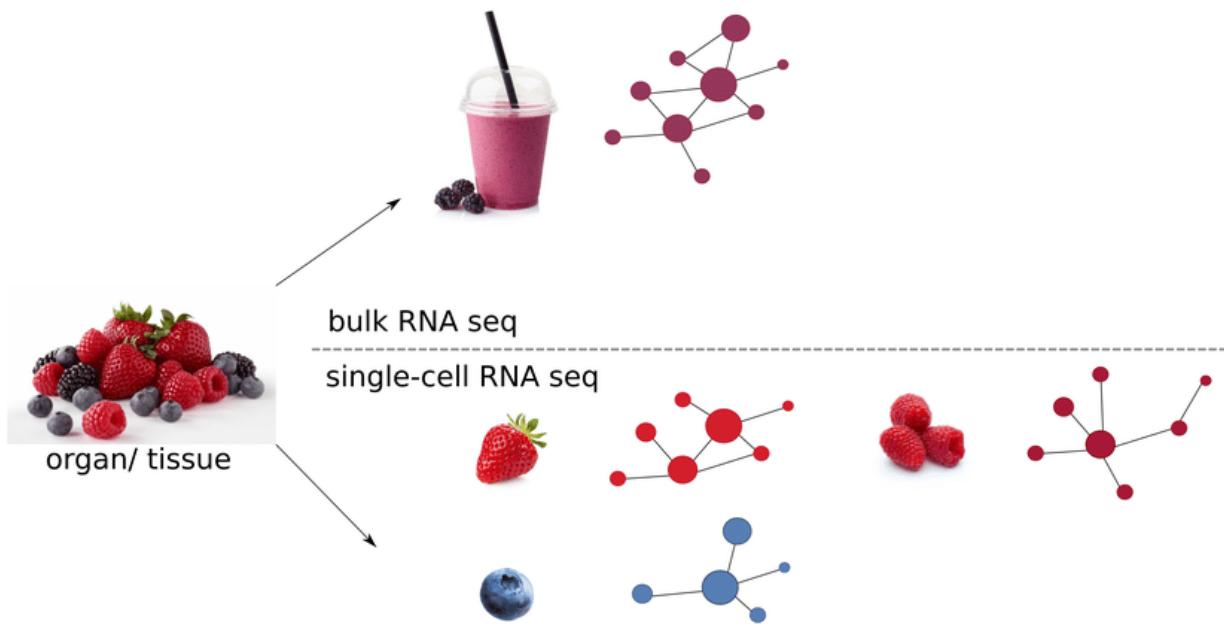


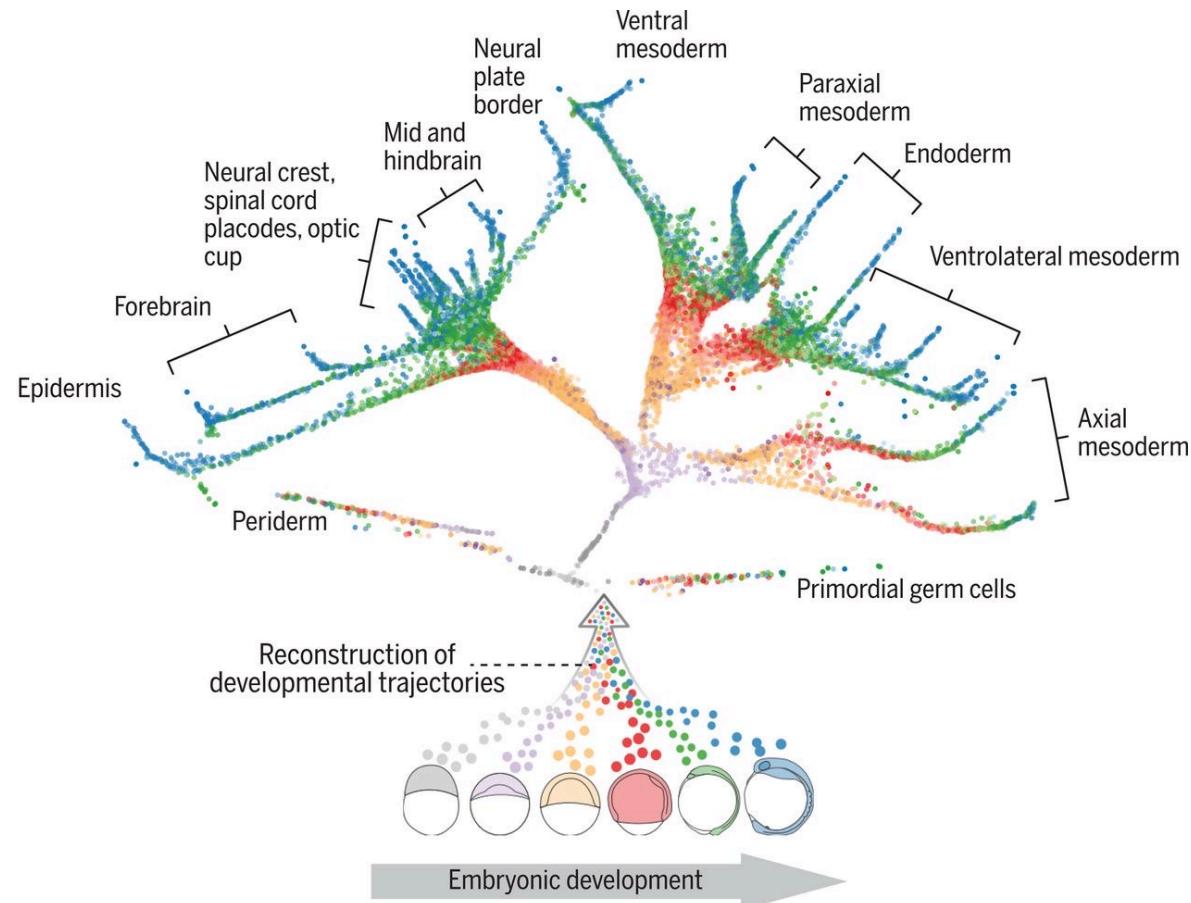
Figure 1. Single cell gene expression profiling reveals cellular heterogeneity that is masked by bulk RNA-seq methods.

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?



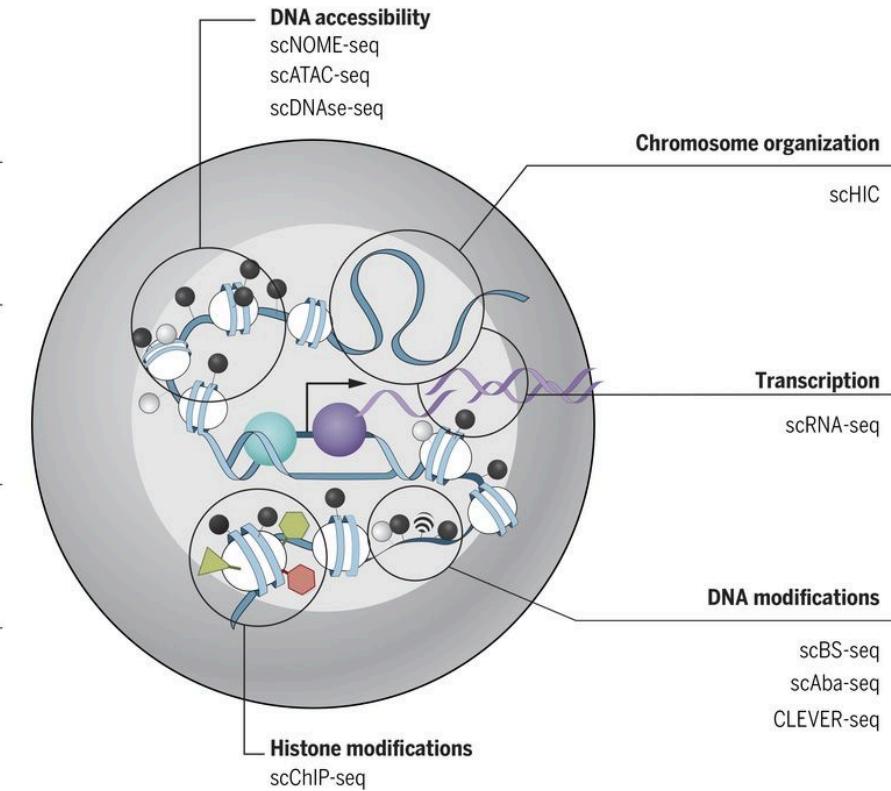
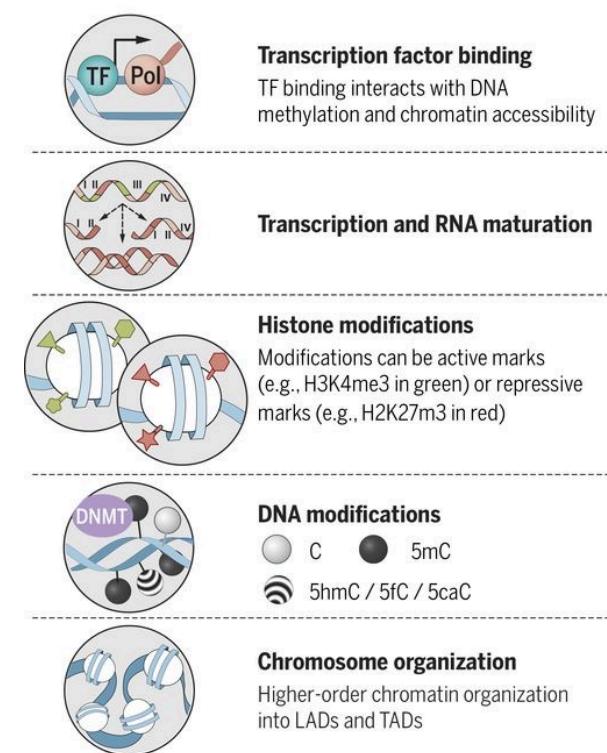
Cell type vs cell lineage vs cell state



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

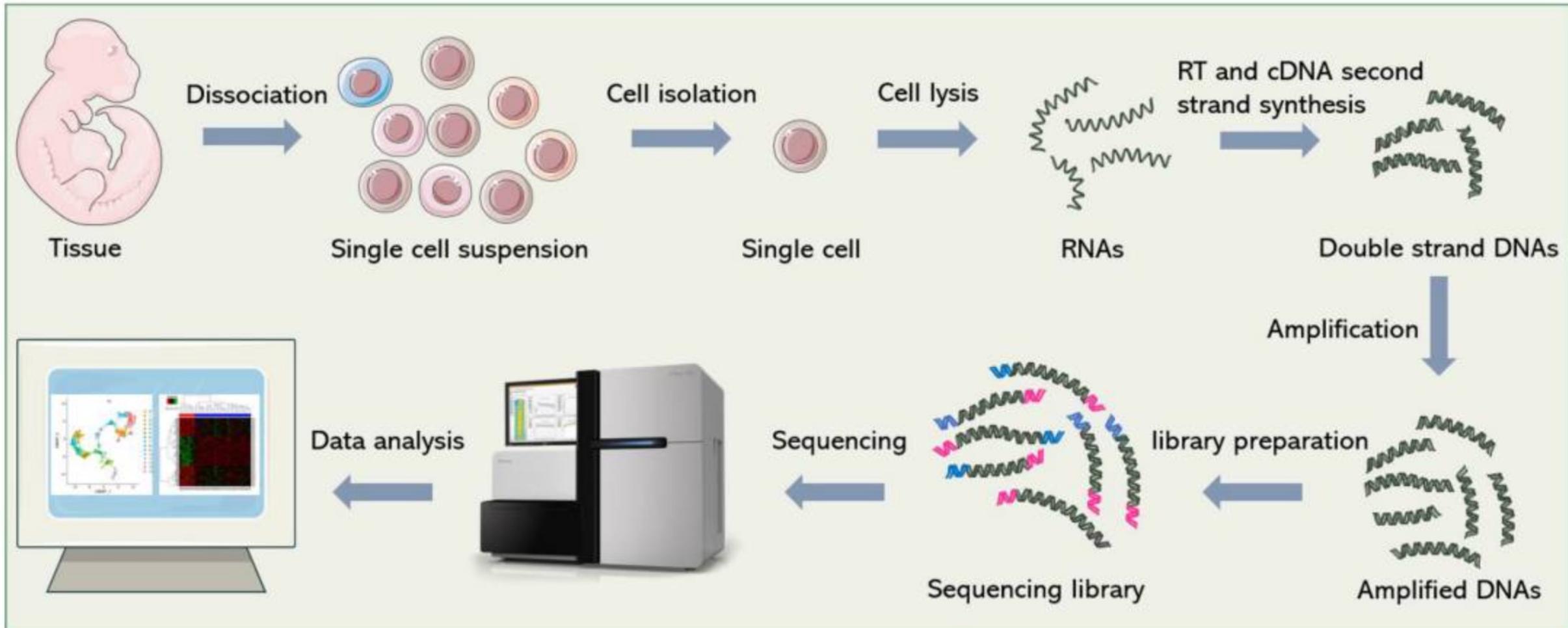
Multimodal single-cell measurements

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

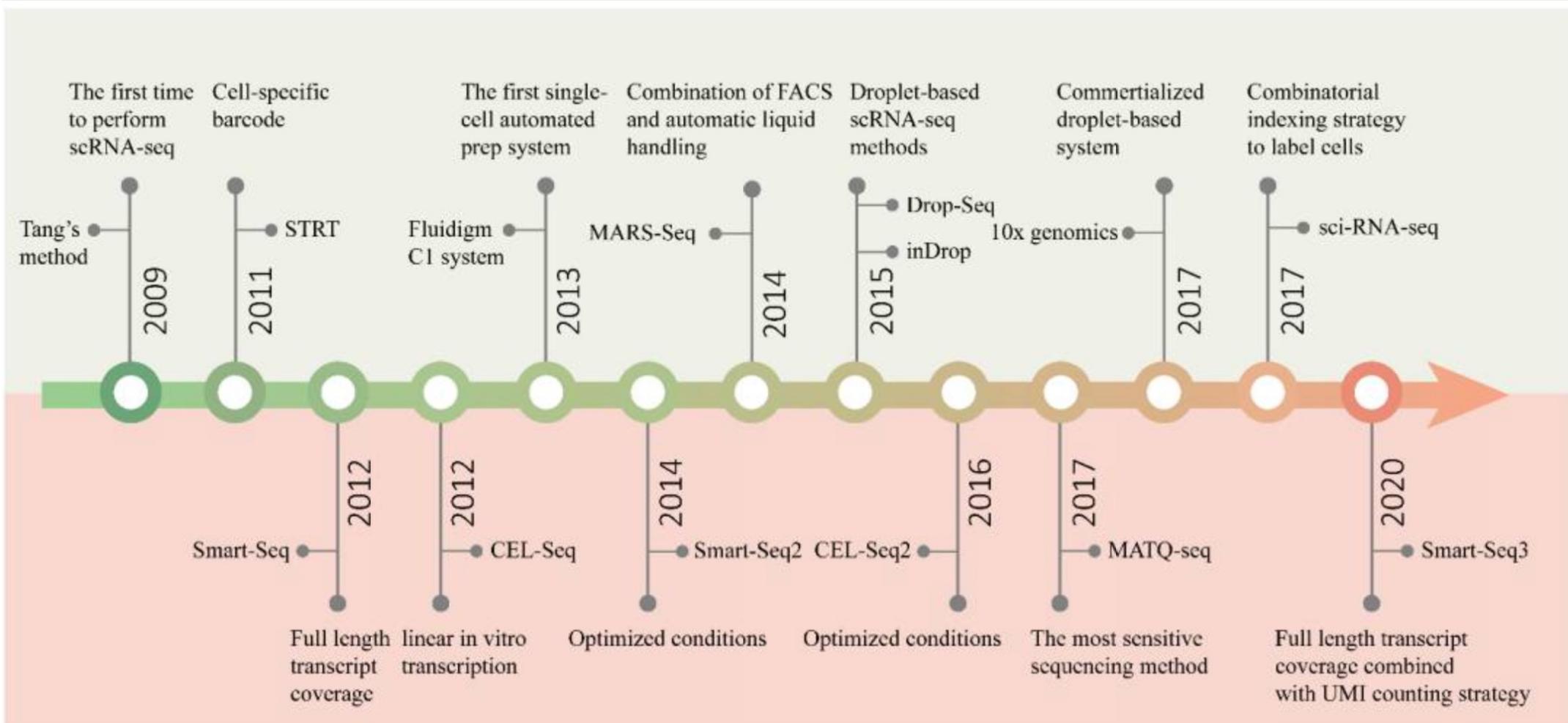


Kelsey et al. Science 2017

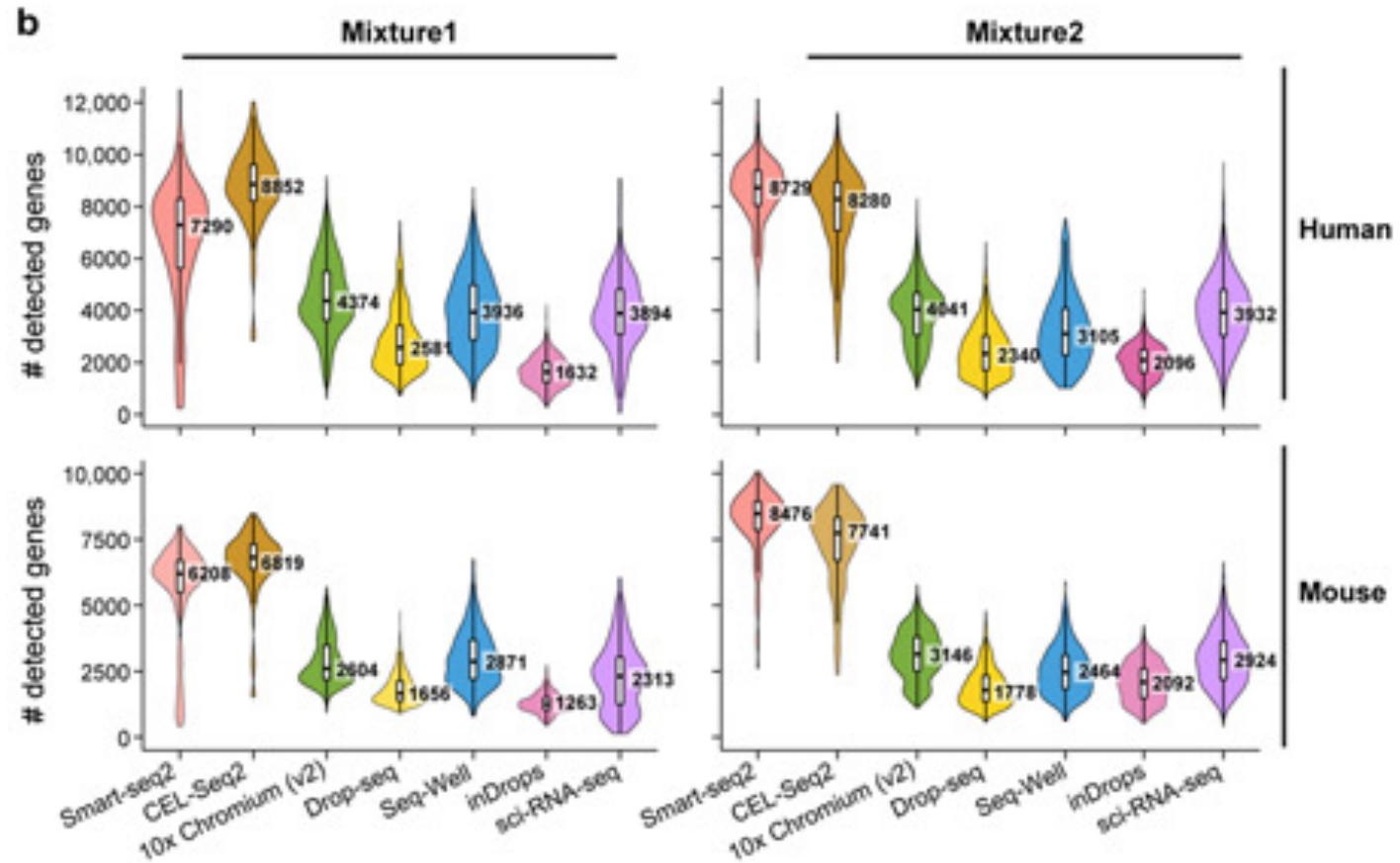
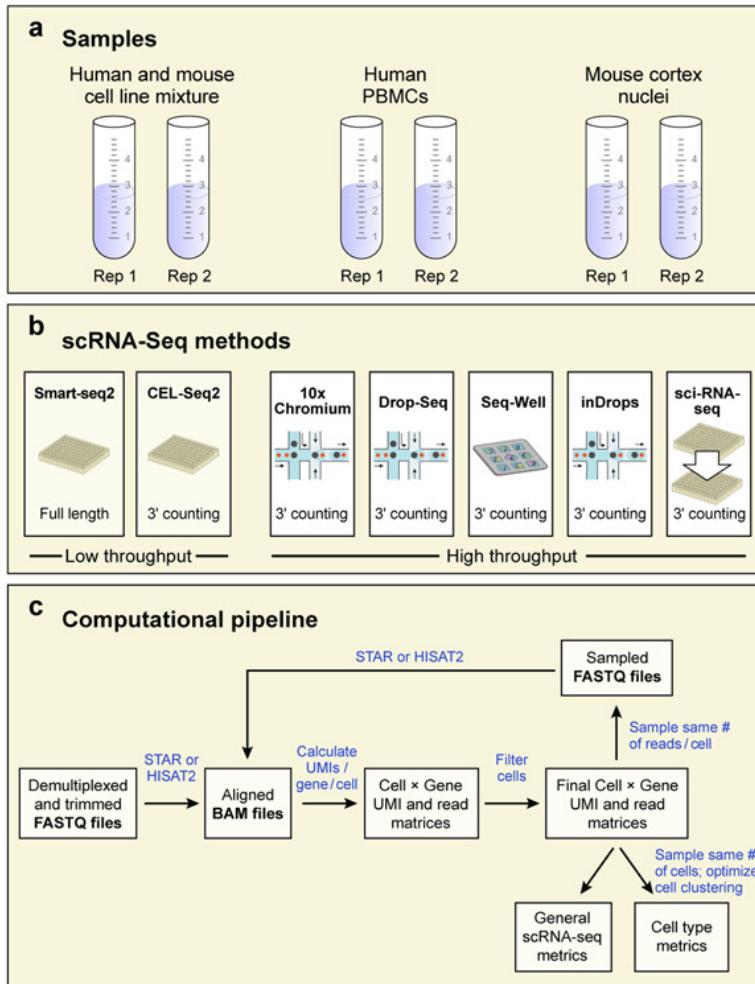
Basic overview of single cell RNA analysis



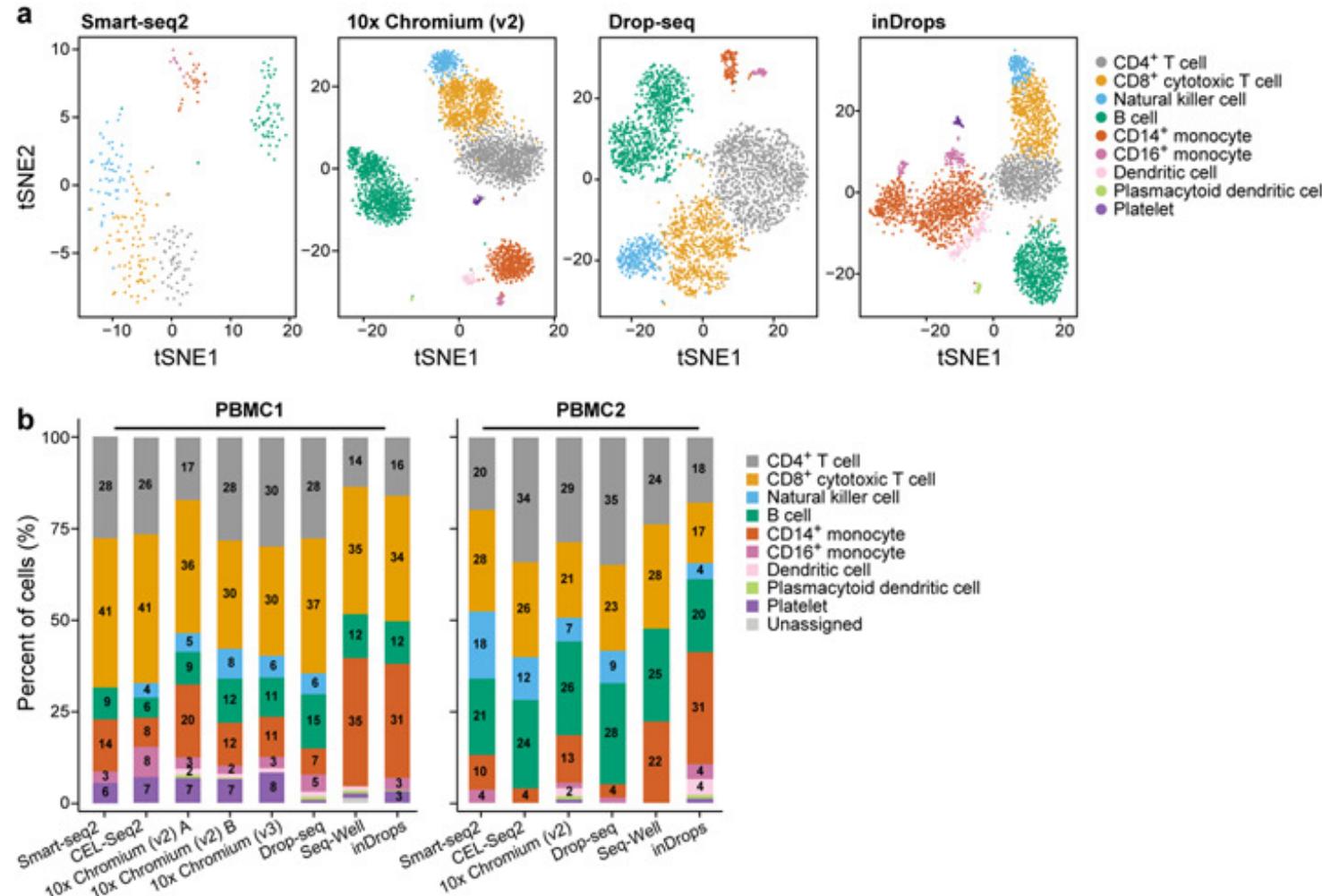
History of scRNA-seq Technology development



Multiple platforms have emerged ... Each with trade-offs...



Multiple platforms have emerged ... Each with trade-offs...

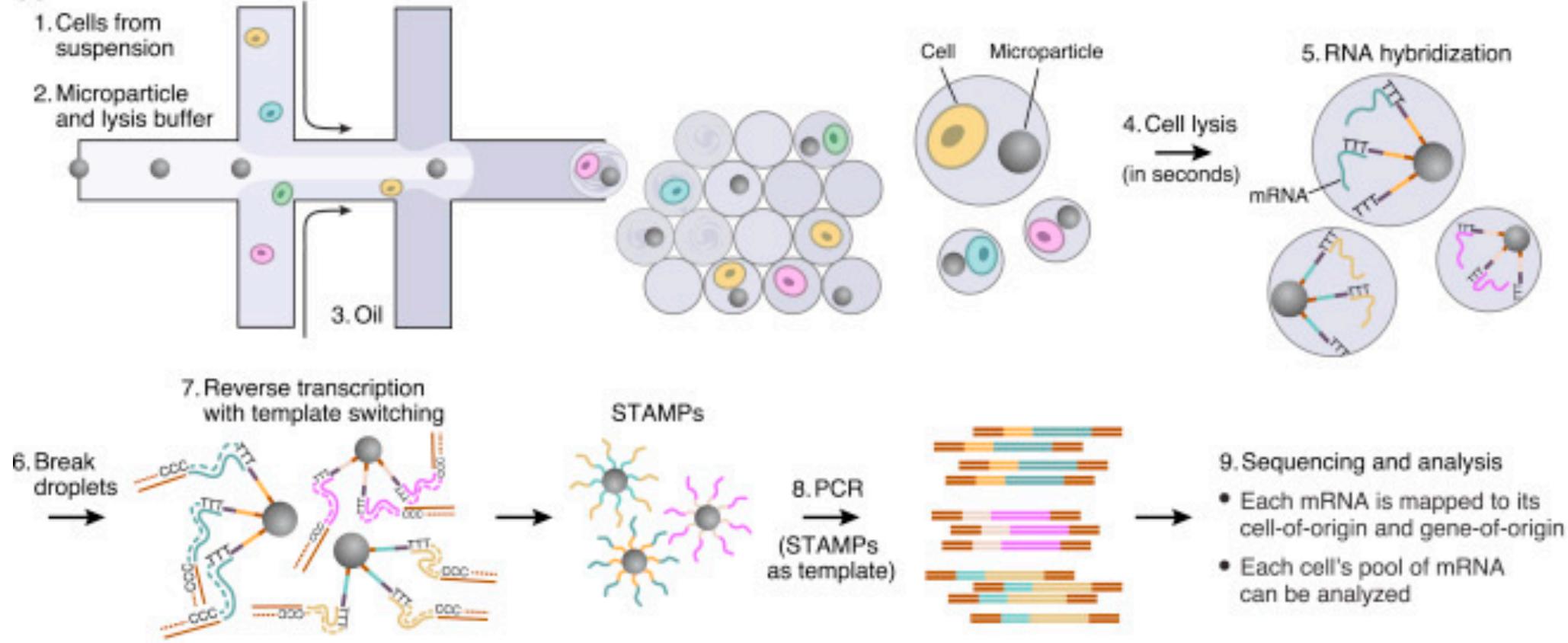


Additional considerations

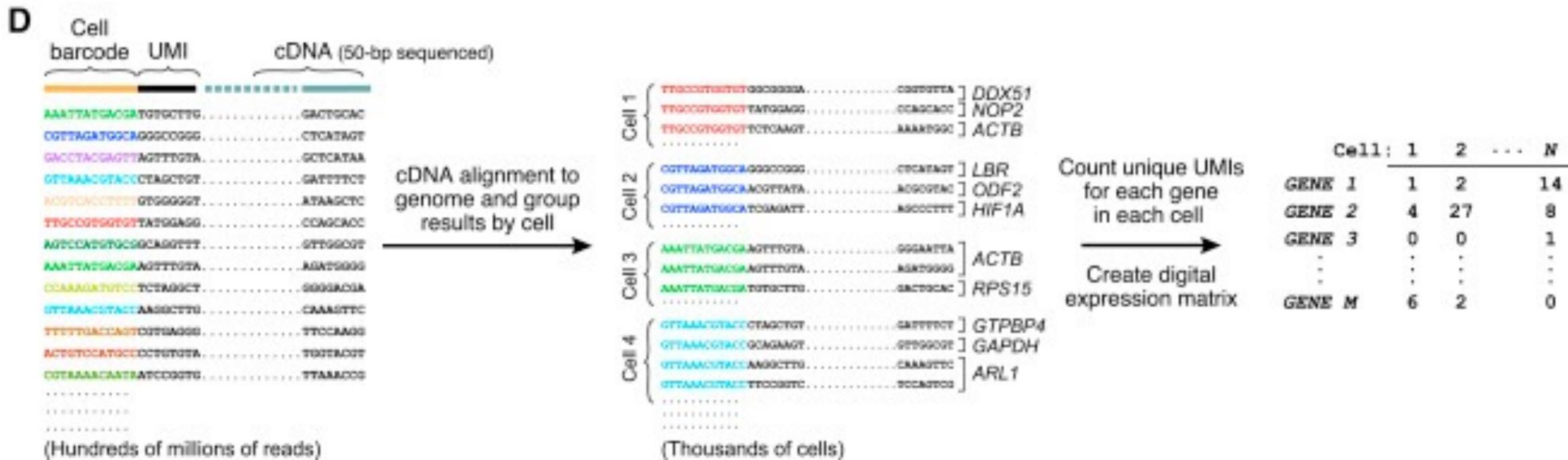
- Size of cells
- Expected % of cell population of interest
- Depth of detection needed for experimental question
- Accessibility

Intro to 10X approach

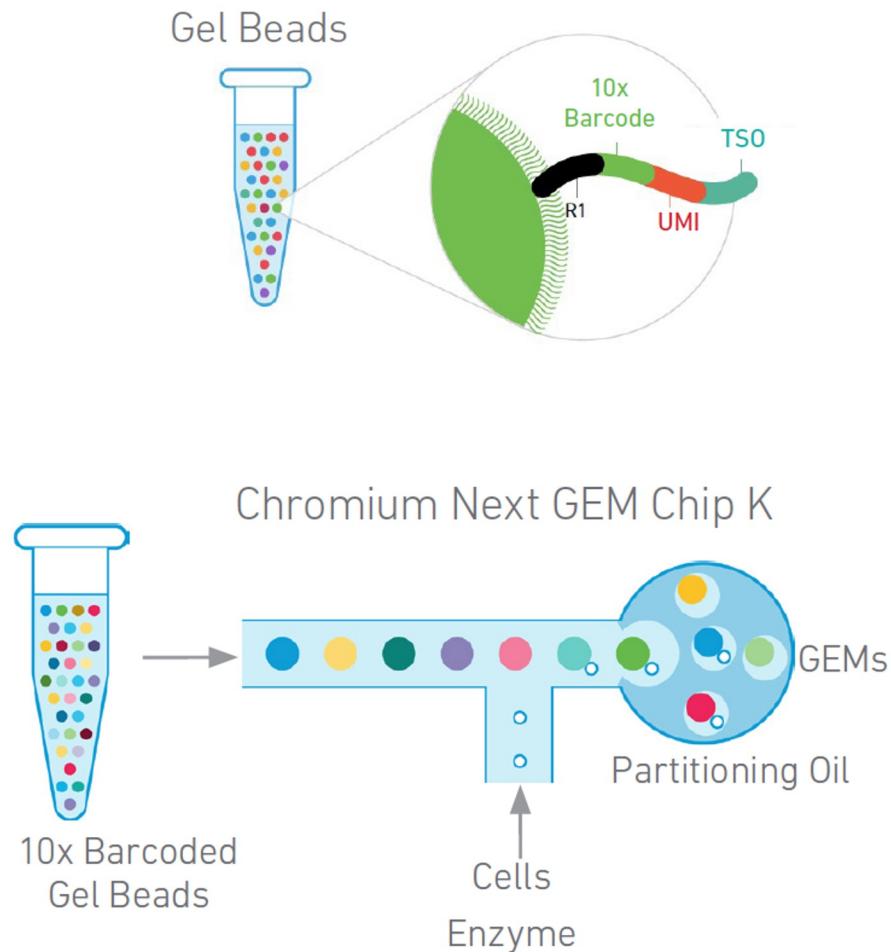
A



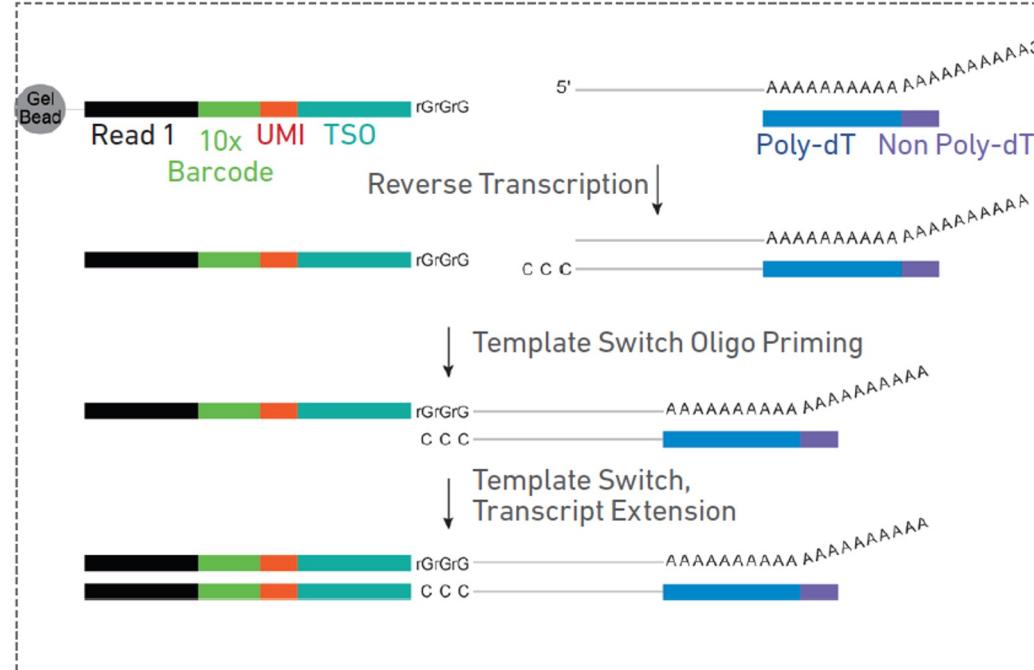
Intro to 10X approach



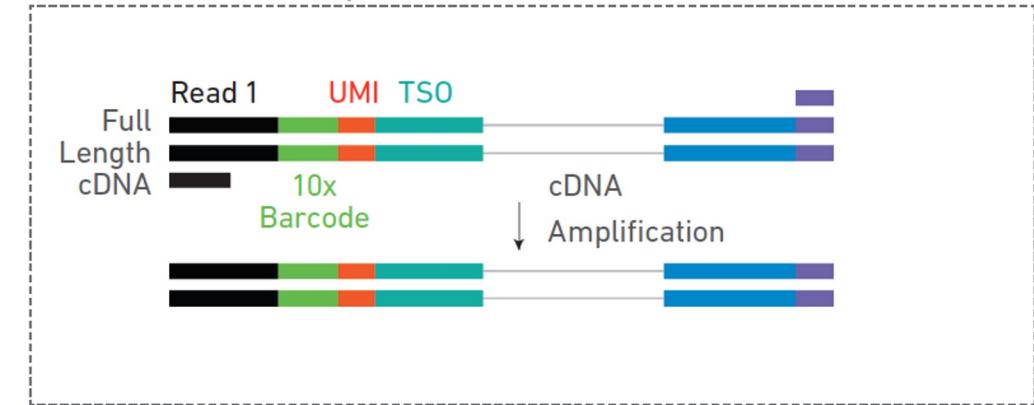
10X (5') v1/v2 Approach



Inside individual GEMs

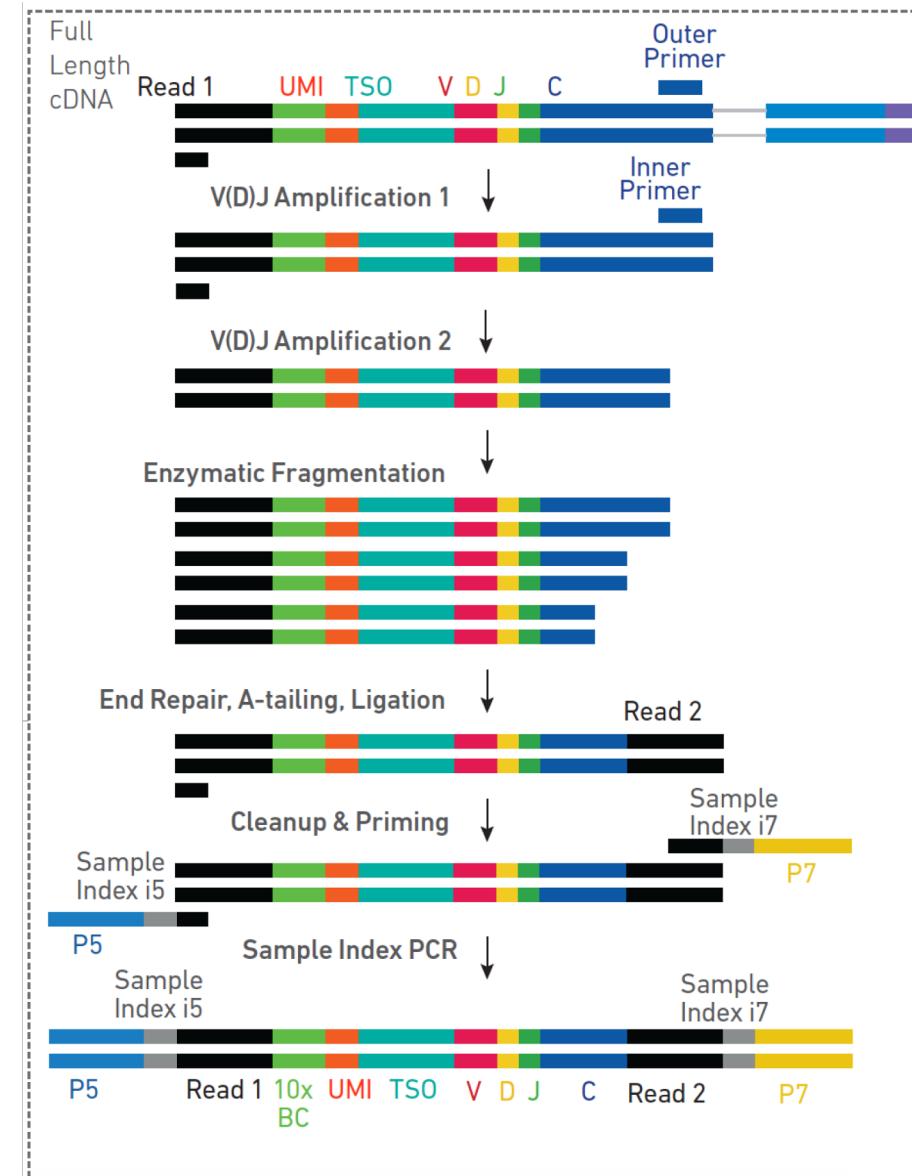
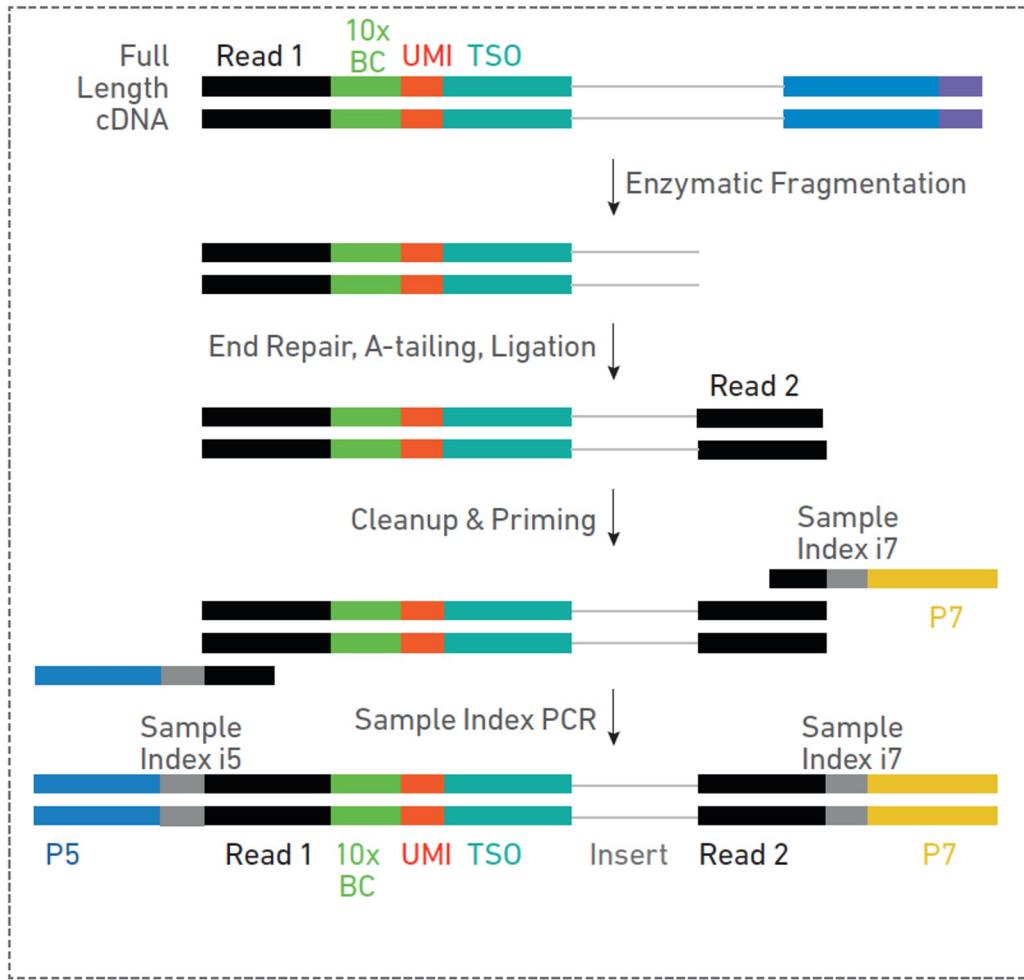


Pooled cDNA amplification



10X (5') v2 Approach

Pooled amplified cDNA processed in bulk

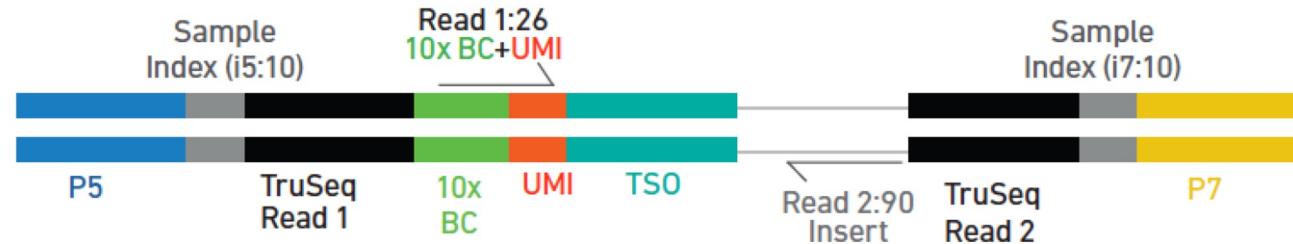


10X Single Cell 5' GEX and V(D)J dual index libraries

Chromium Single Cell V(D)J Dual Index Library

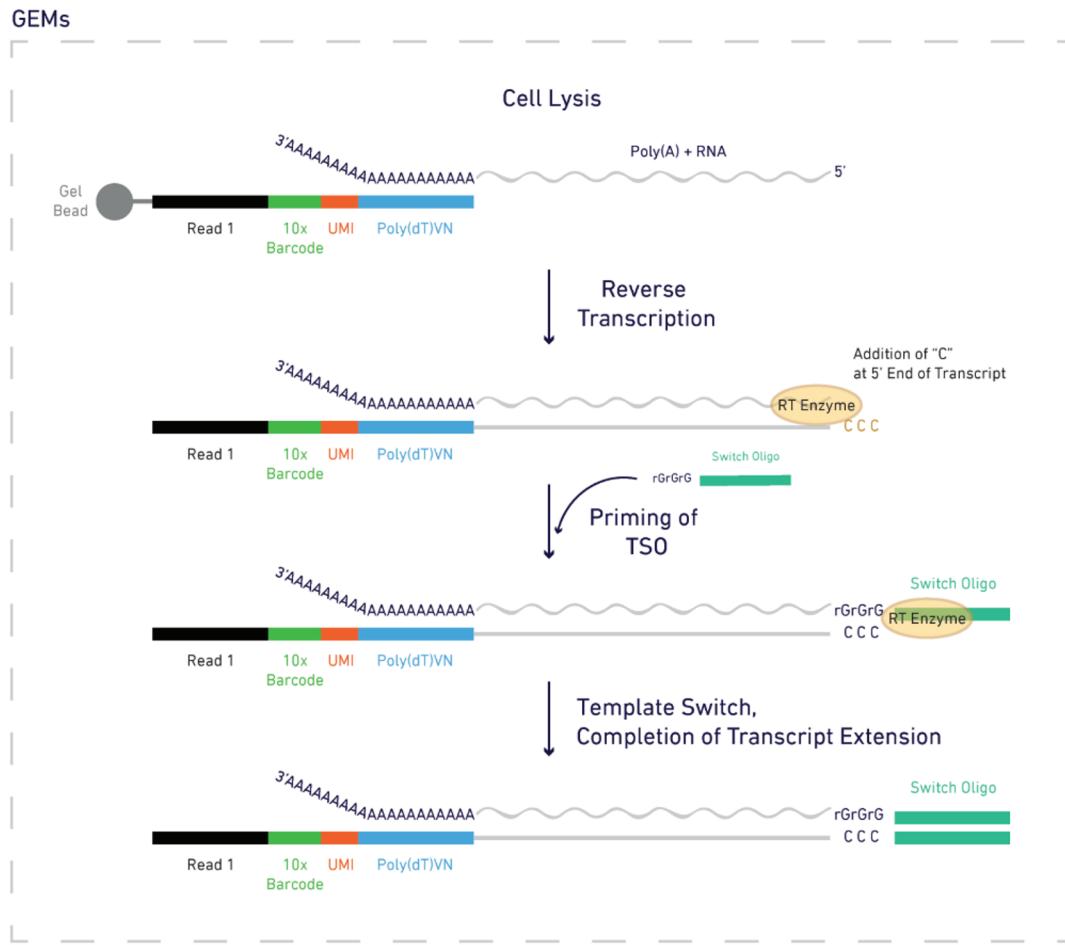


Chromium Single Cell 5' Gene Expression Dual Index Library

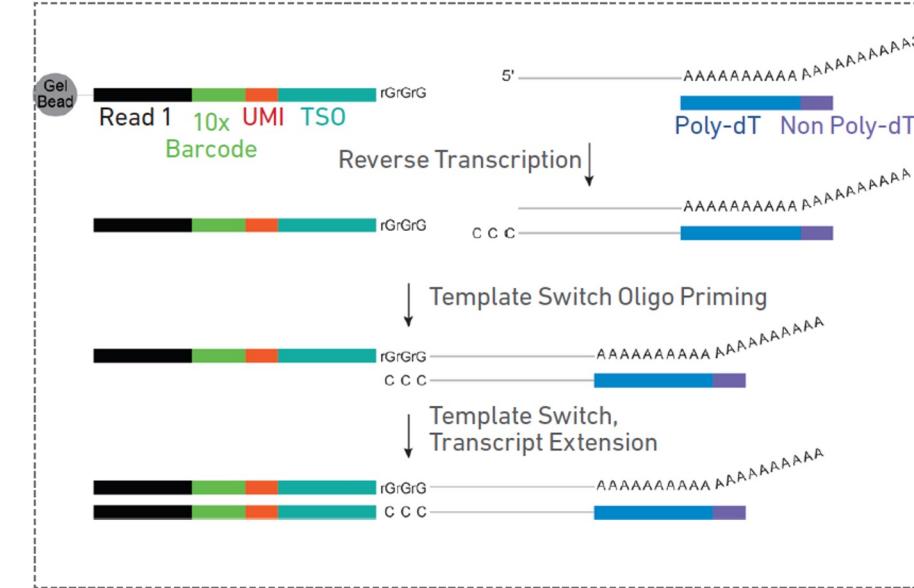


5' vs 3' approach - polyT and TSO switch places

Inside individual GEMs

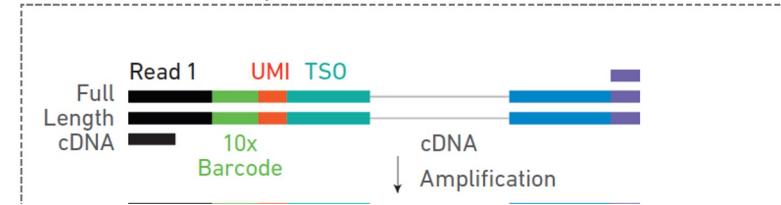


Inside individual GEMs

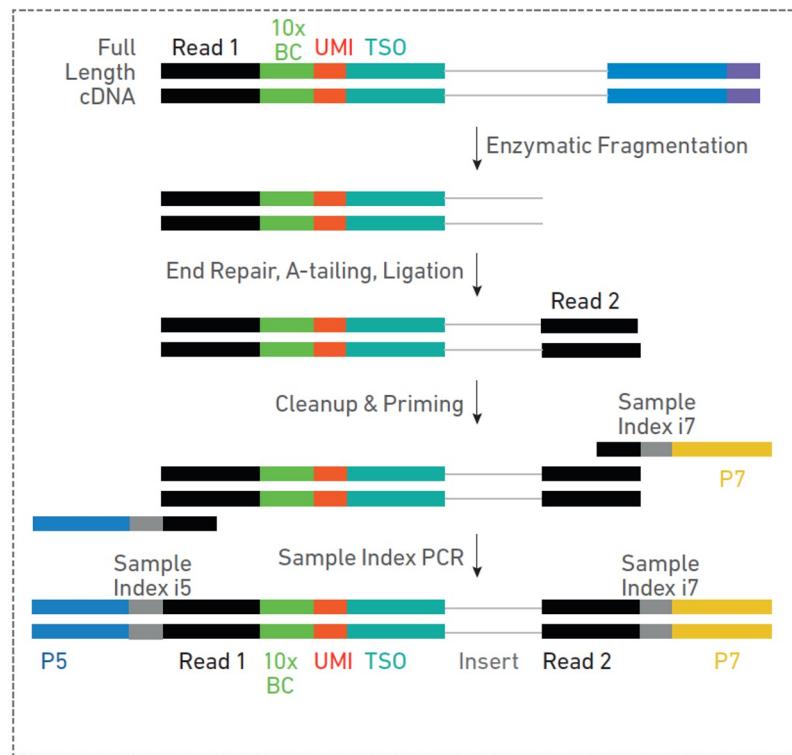


5' vs 3' approach - Fragmentation and PCR steps result in different ends of transcripts being kept

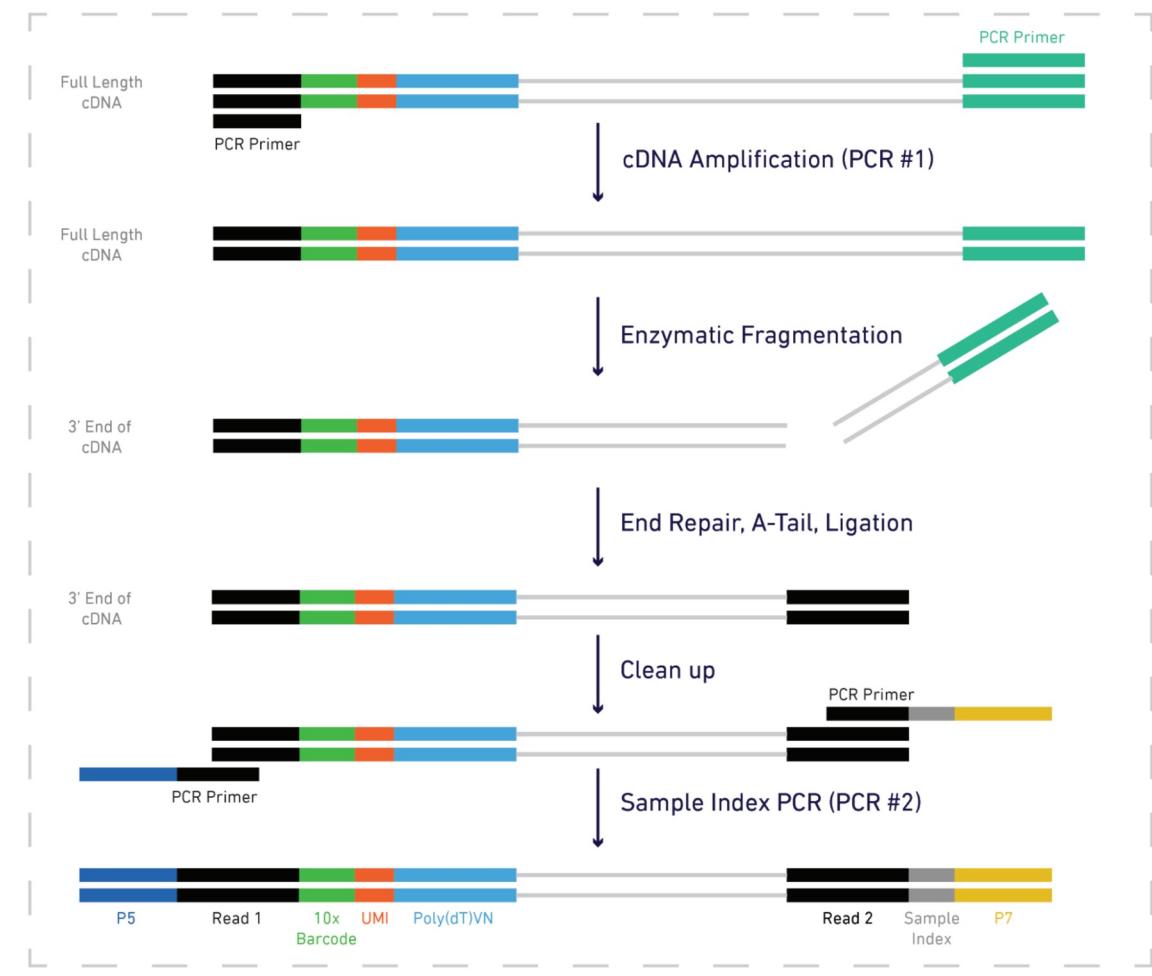
Pooled cDNA amplification



Pooled amplified cDNA processed in bulk



5' -> 3'

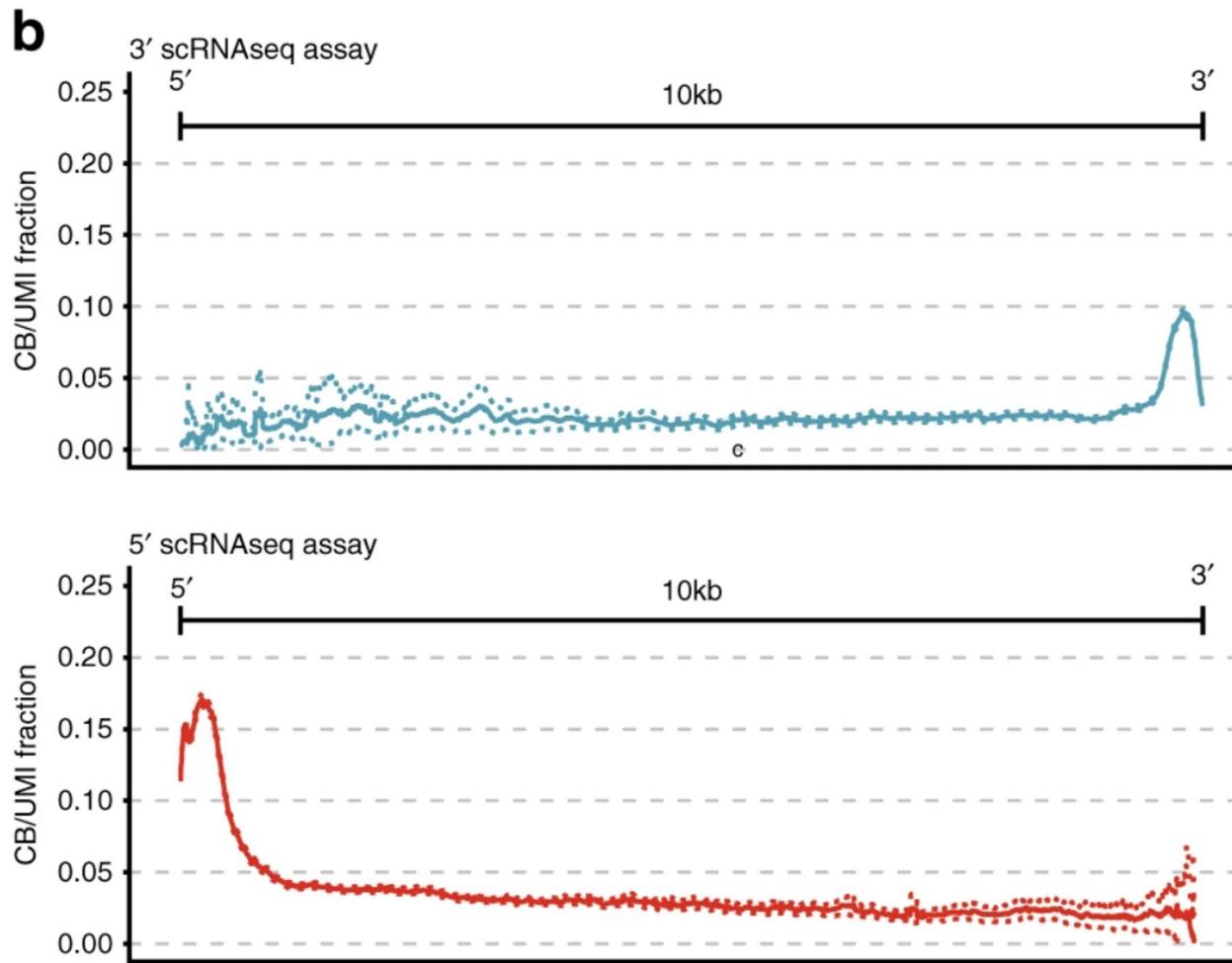


19

3' <- 5'

rnabio.org

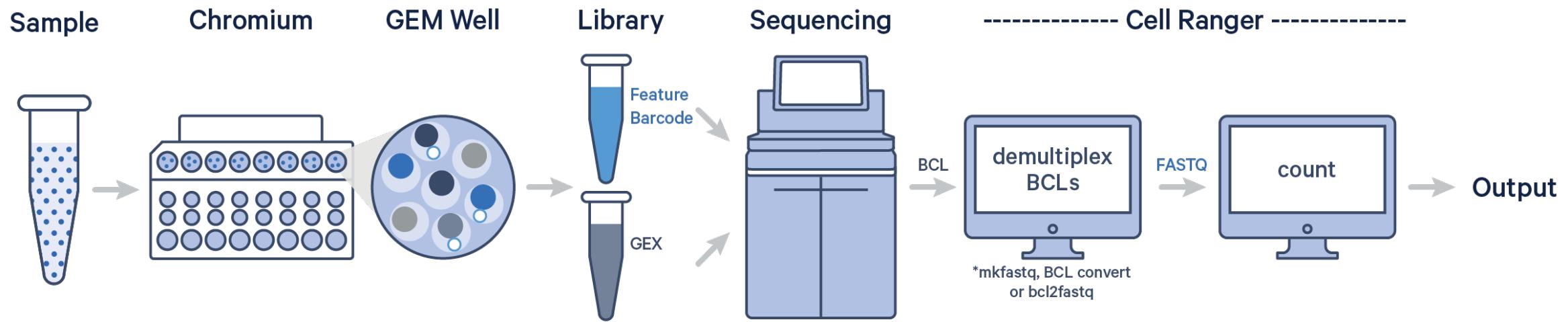
5' vs 3' approach results in different end bias



5' vs 3' 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

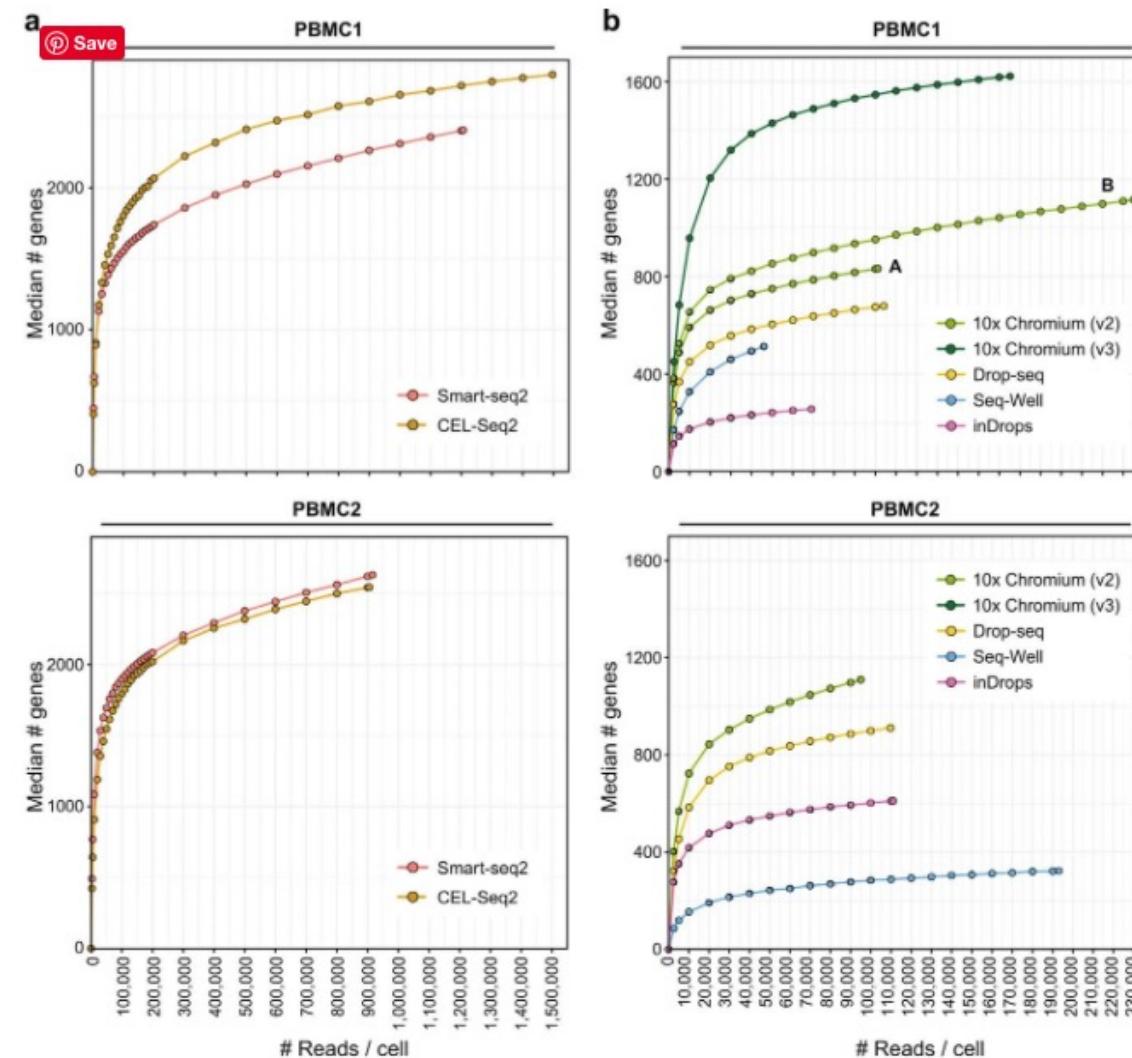


How deeply do you need to sequence?

General Rule:
Achieve 90% saturation

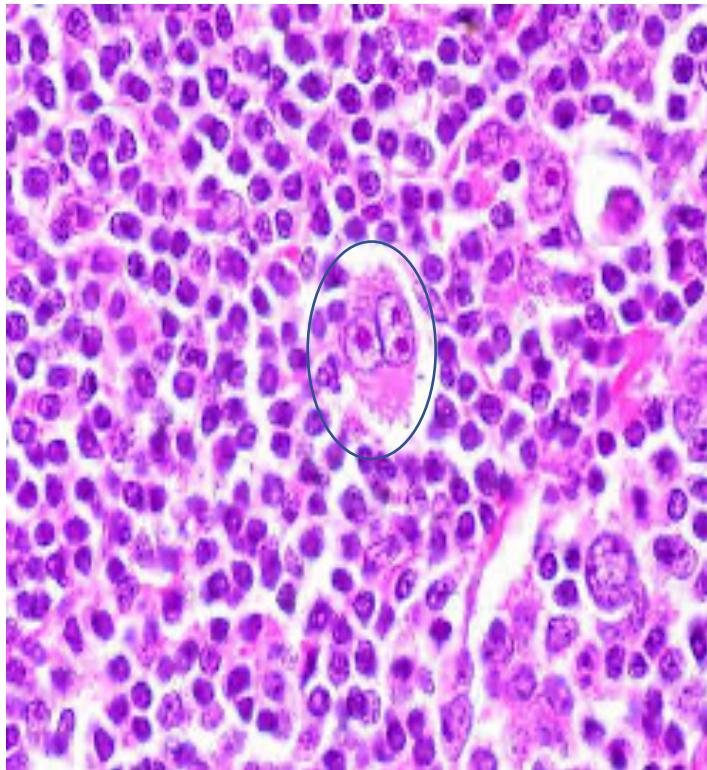
Official Recommendations (reads/cell):

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



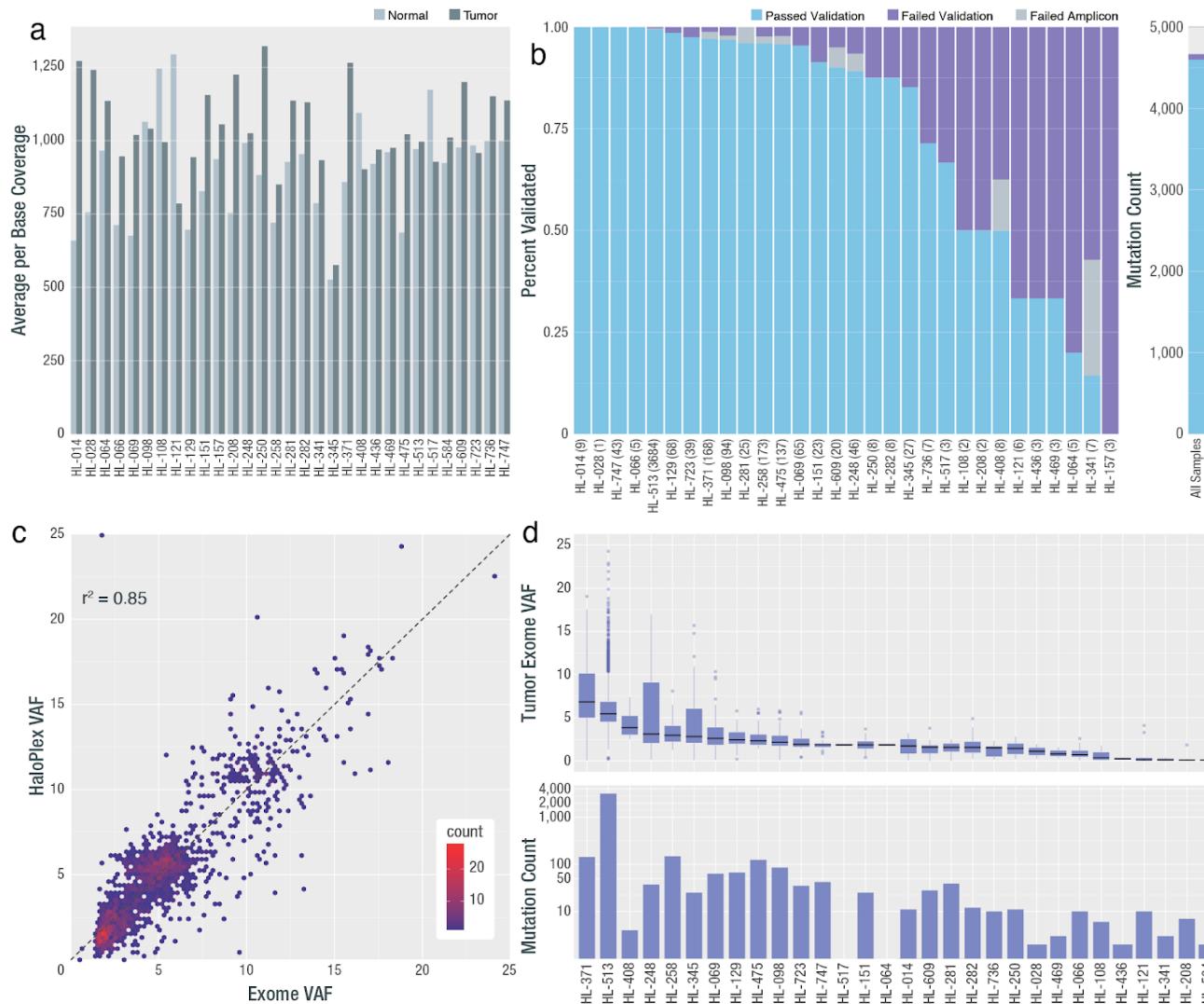
Felicia Gomez



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

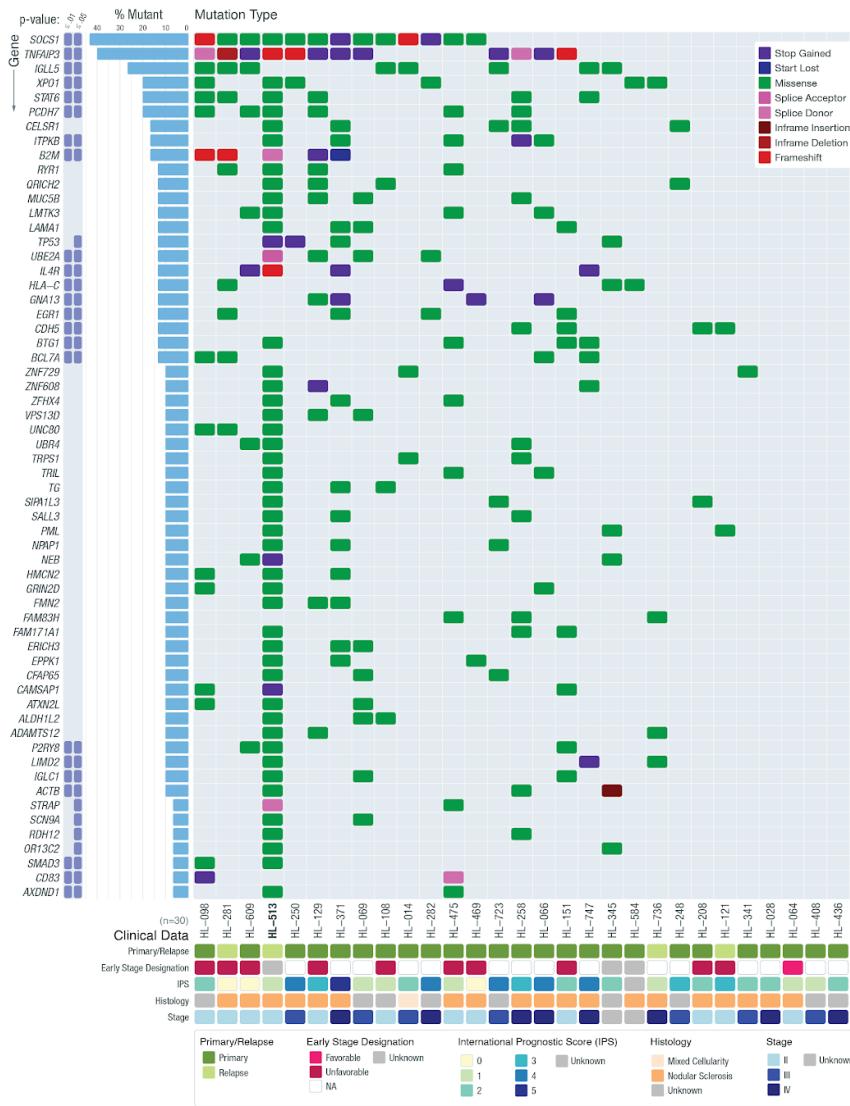
- IDT Exome capture reagent
- Three KAPA libraries were constructed/sample
- Libraries were sequenced across eight lanes of an Illumina HiSeq
- Somatic Variant Calling
 - SNV were called using 5 variant callers
 - Indels were called using 4 variant callers
- Target depth ~1000x
- Validation with targeted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

High coverage and concordance achieved

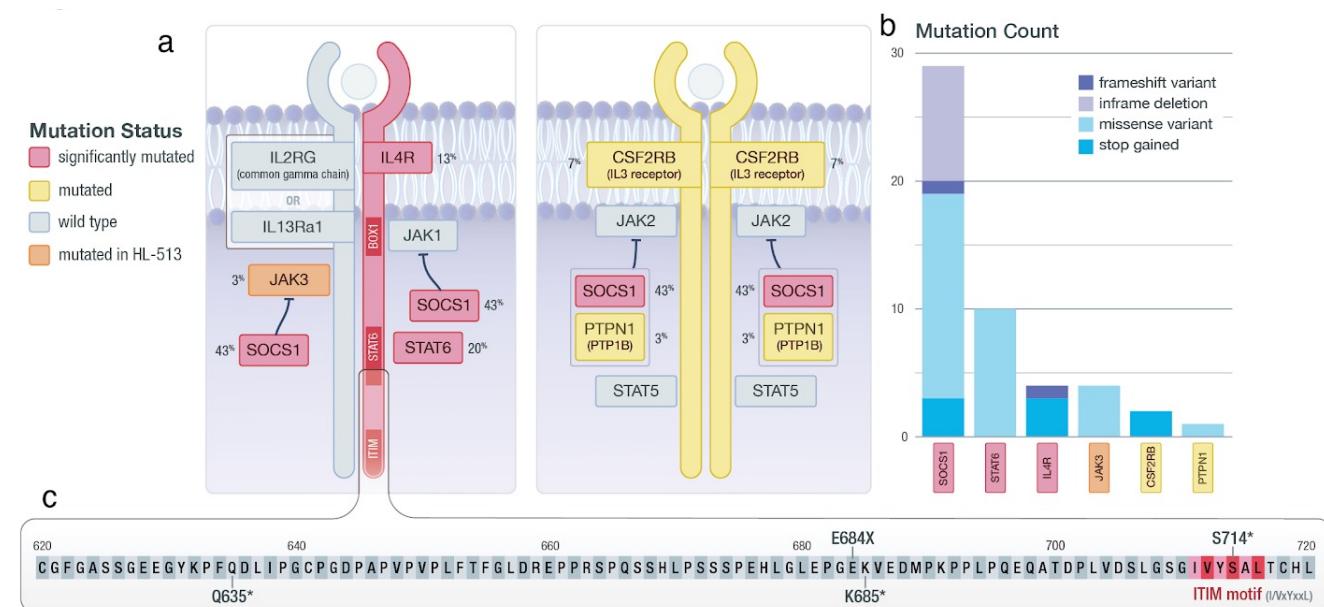


- ~1000X coverage achieved
- High VAF concordance between WES and haloplex validation
- Overall variant validation rate >90%
- Mean mutation burden = 33 protein-coding variants
 - One hypermutator 3,160 variants

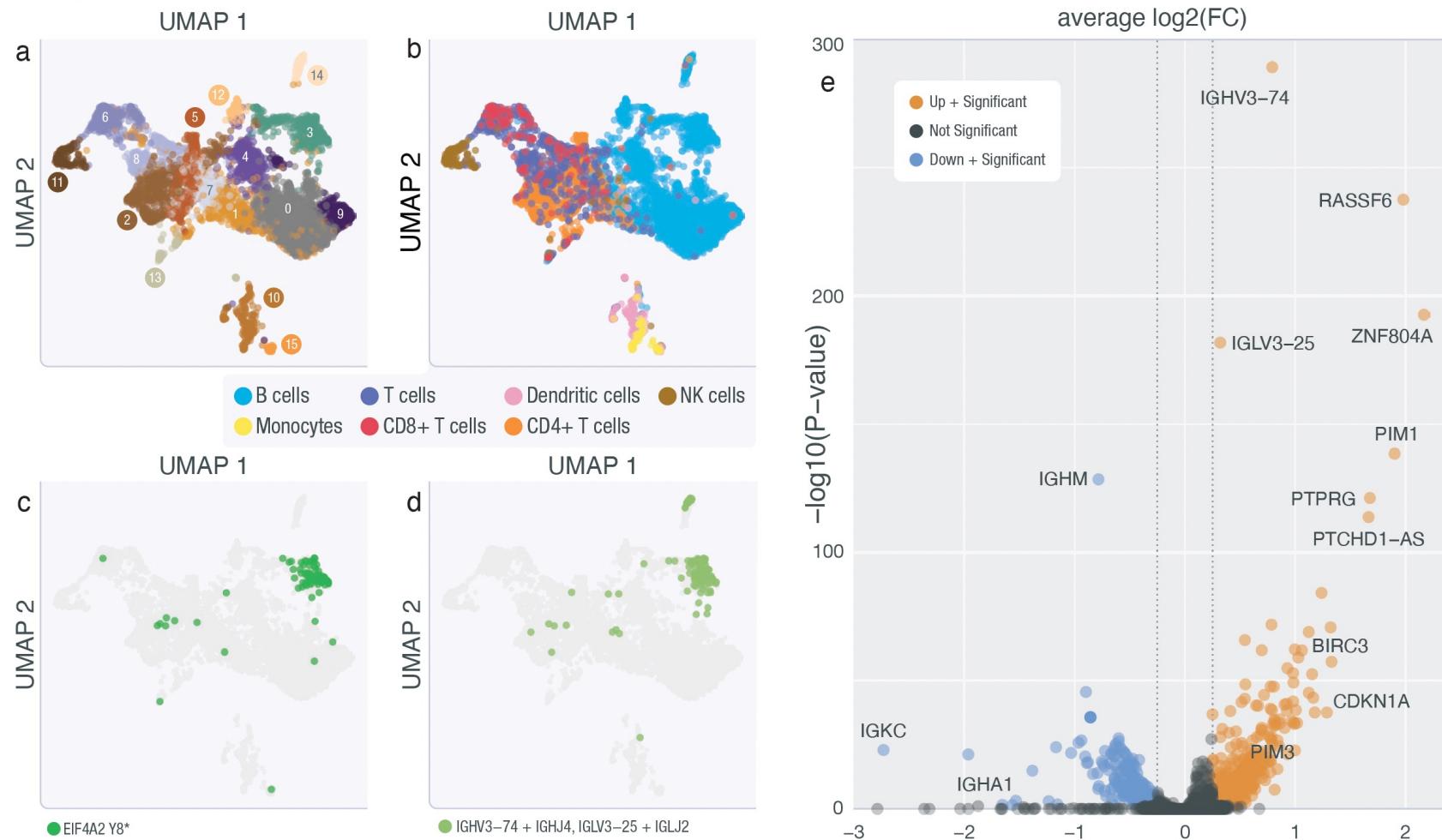
Represents one of the few comprehensive surveys of HL - mutation landscape reveals known and novel genes/pathways/hotspots



- ~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



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. McMichael², Matthew Mosior^{1,2}, Jennifer A. Foltz¹, Zachary L. Skidmore^{1,2}, Eric J. Duncavage⁴, Christopher A. Miller^{1,2}, Haley Abel^{1,2}, Yi-Shan Li⁴, David A. Russler-Germain¹, Kilannin Krysiak^{1,2,3,4}, Marcus P. Watkins¹, Cody A. Ramirez^{1,2}, Alina Schmidt^{1,2}, Fernanda Martins Rodrigues^{1,2}, Lee Trani², Ajay Khanna¹, Julia A. Wagner¹, Robert S. Fulton², Catrina C. Fronick², Michelle D. O'Laughlin², Timothy Schappe¹, Amanda F. Cashen¹, Neha Mehta-Shah¹, Brad S. Kahl¹, Jason Walker², Nancy L. Bartlett¹, Malachi Griffith^{1,2,3,5}, Todd A. Fehniger^{1,3}, and Obi L. Griffith^{1,2,3,5}

We are on a Coffee Break & Networking Session