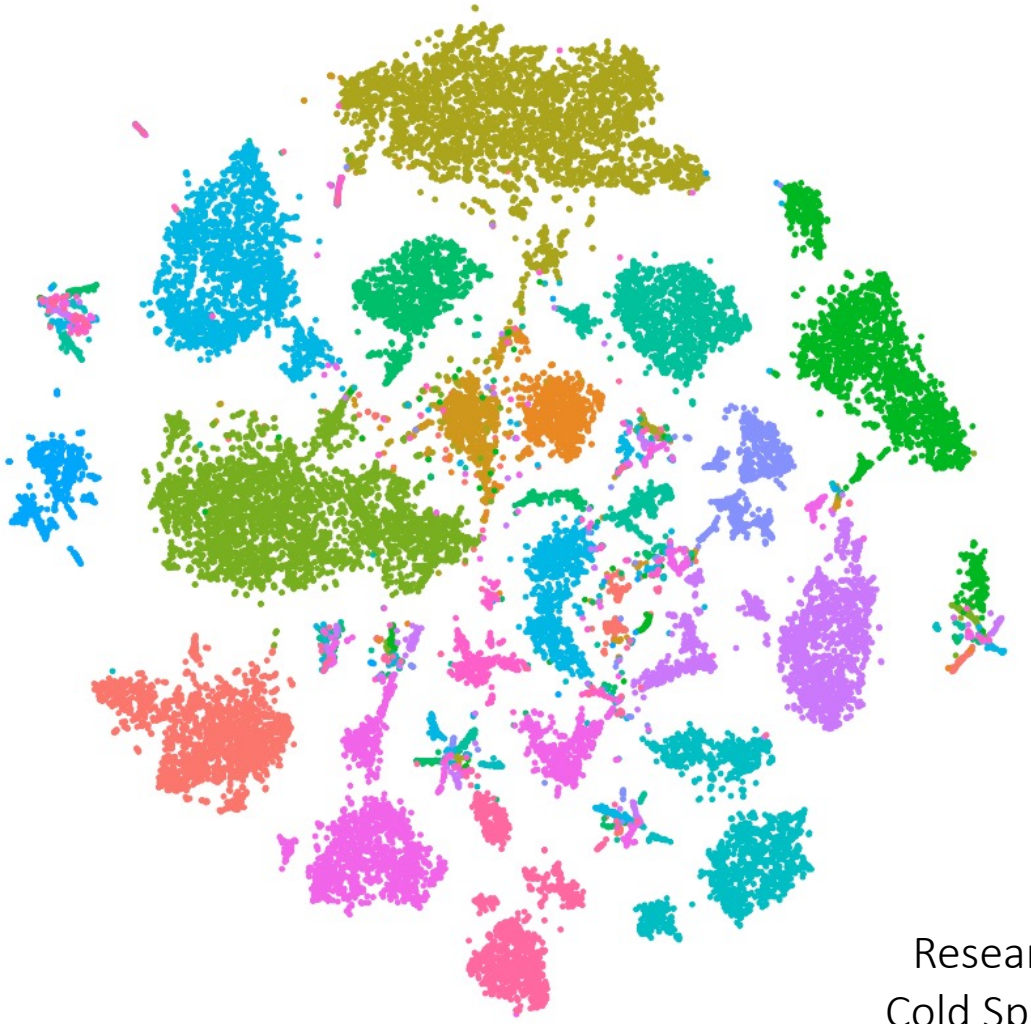
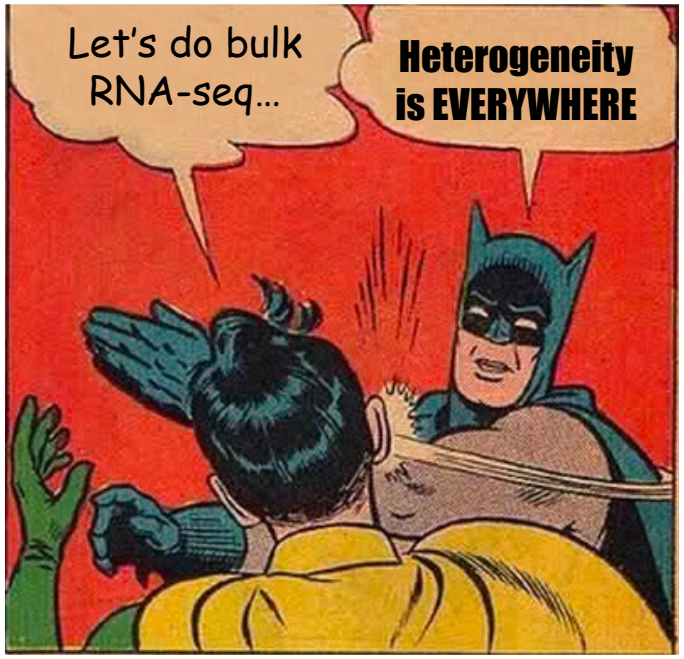
Single Cell Sequencing

CSHL Course: Advanced Sequencing Technologies &Applications

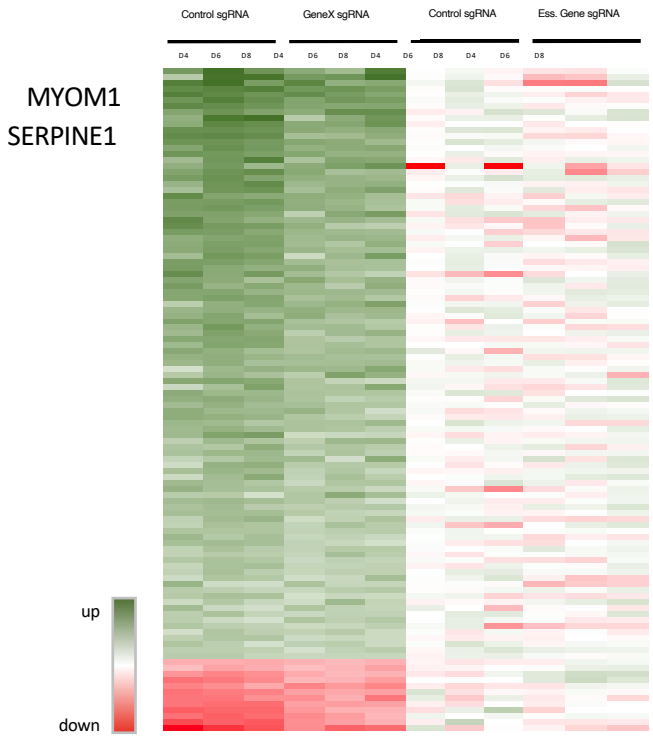


Jon Preall
Research Associate Professor
Cold Spring Harbor Laboratory

Why Sequence Single Cells?



The Importance of Single Cell Resolution

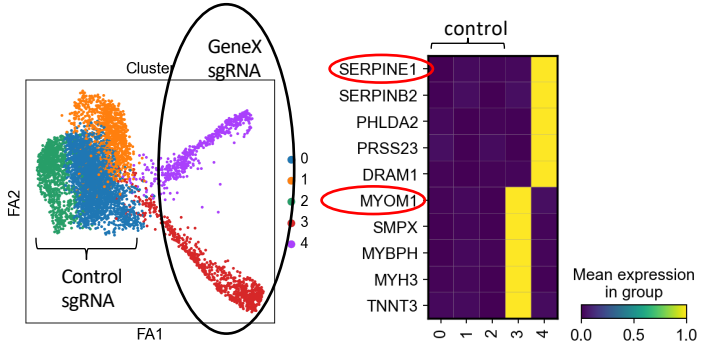


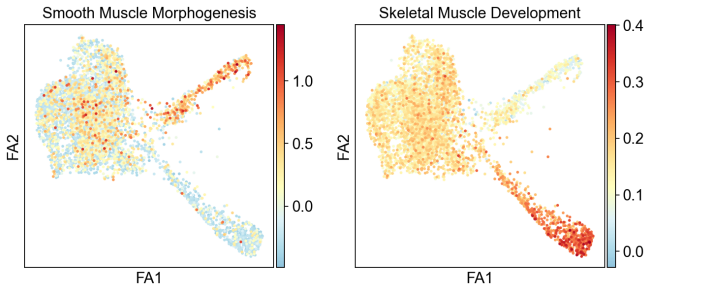
Human Rhabdomyosarcoma cell line

CRISPR KO Driver gene

Bulk RNAseq

10X scRNAseq



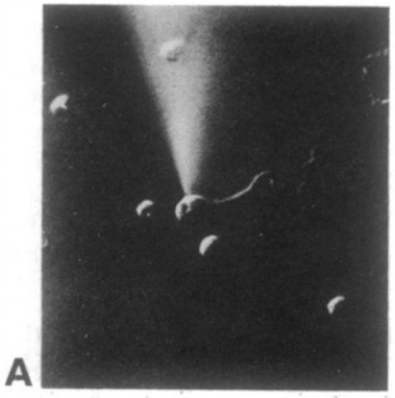


Analysis of gene expression in single live neurons

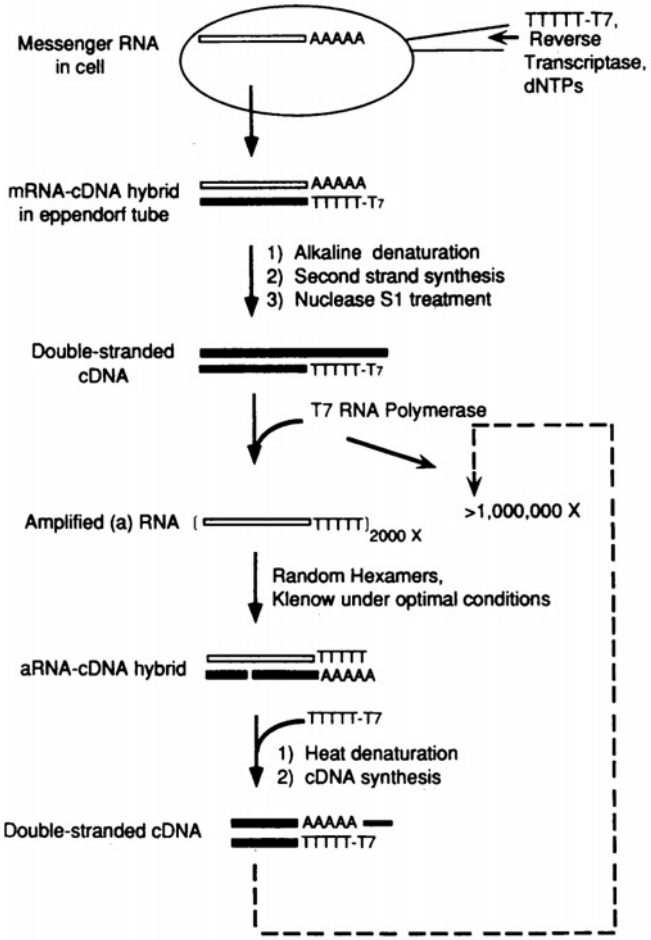
(amplified, antisense RNA/expression profile/mRNA complexity/pyramidal cell)

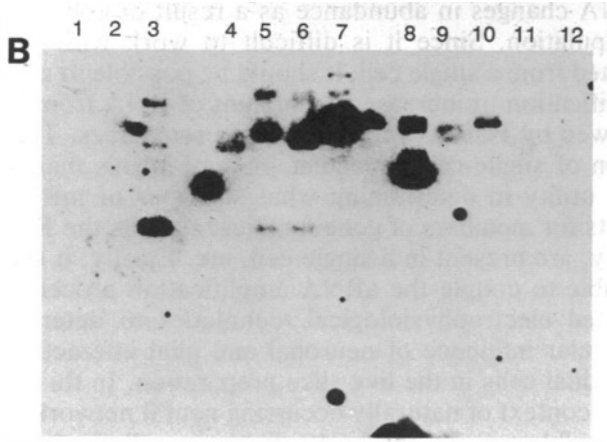
JAMES EBERWINE*†‡, HERMES YEH§, KEVIN MIYASHIRO*, YANXIANG CAO*, SURESH NAIR*, RICHARD FINNELL*¶, MARTHA ZETTEL§, AND PAUL COLEMAN§

Departments of *Pharmacology and †Psychiatry, University of Pennsylvania Medical School, Philadelphia, PA 19104; and Department of §Neurobiology and Anatomy, University of Rochester Medical Center, Rochester, NY 14642



Microinjection of cDNA synthesis reagents directly into single neurons





Southern Blot
Plasmid standards containing gene of interest
Probed with aRNA

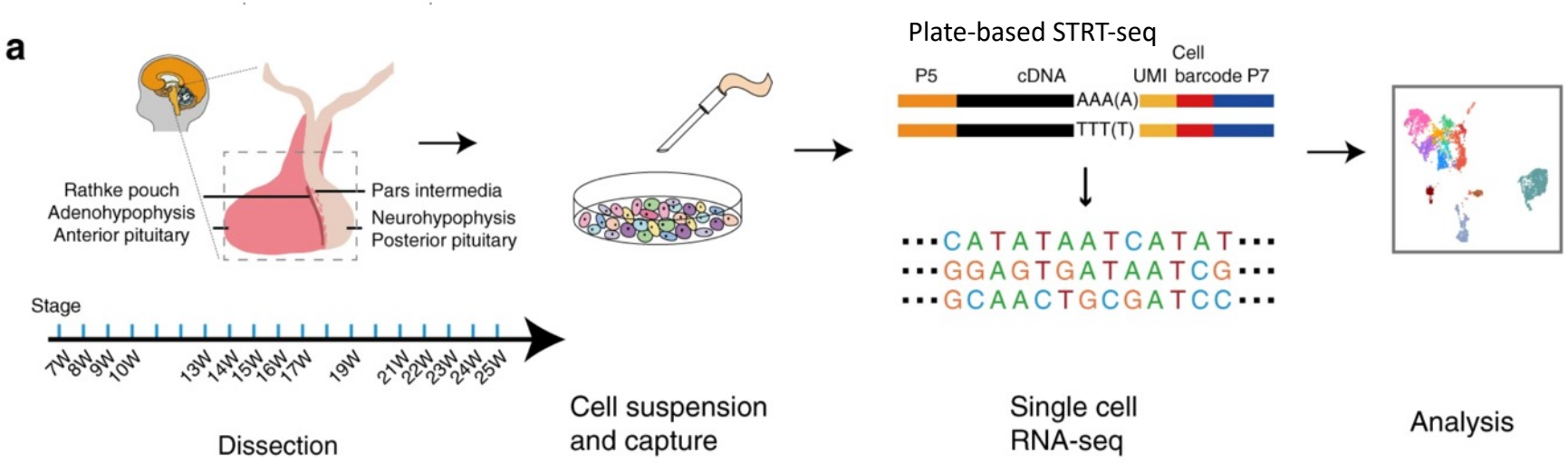
1992

Article | Open Access | Published: 19 October 2020

Single-cell transcriptomics identifies divergent developmental lineage trajectories during human pituitary development

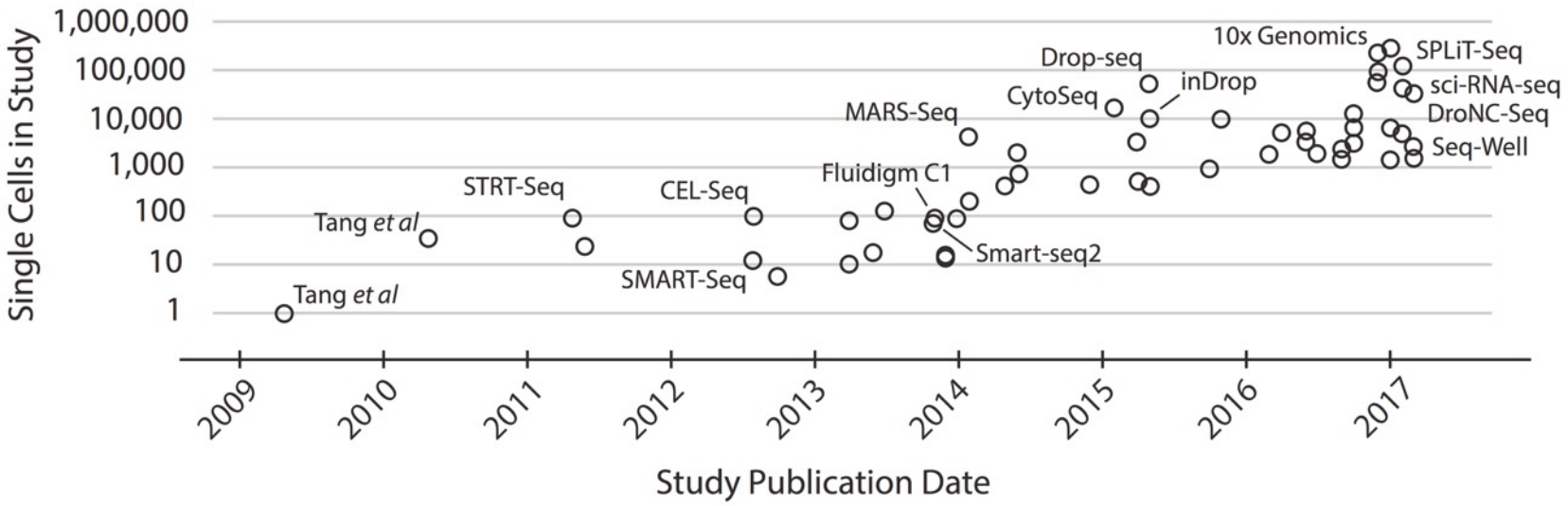
Shu Zhang, Yueli Cui, Xinyi Ma, Jun Yong, Liying Yan, Ming Yang, Jie Ren, Fuchou Tang, Lu Wen ✉ & Jie Qiao ✉

Nature Communications **11**, Article number: 5275 (2020) | Cite this article



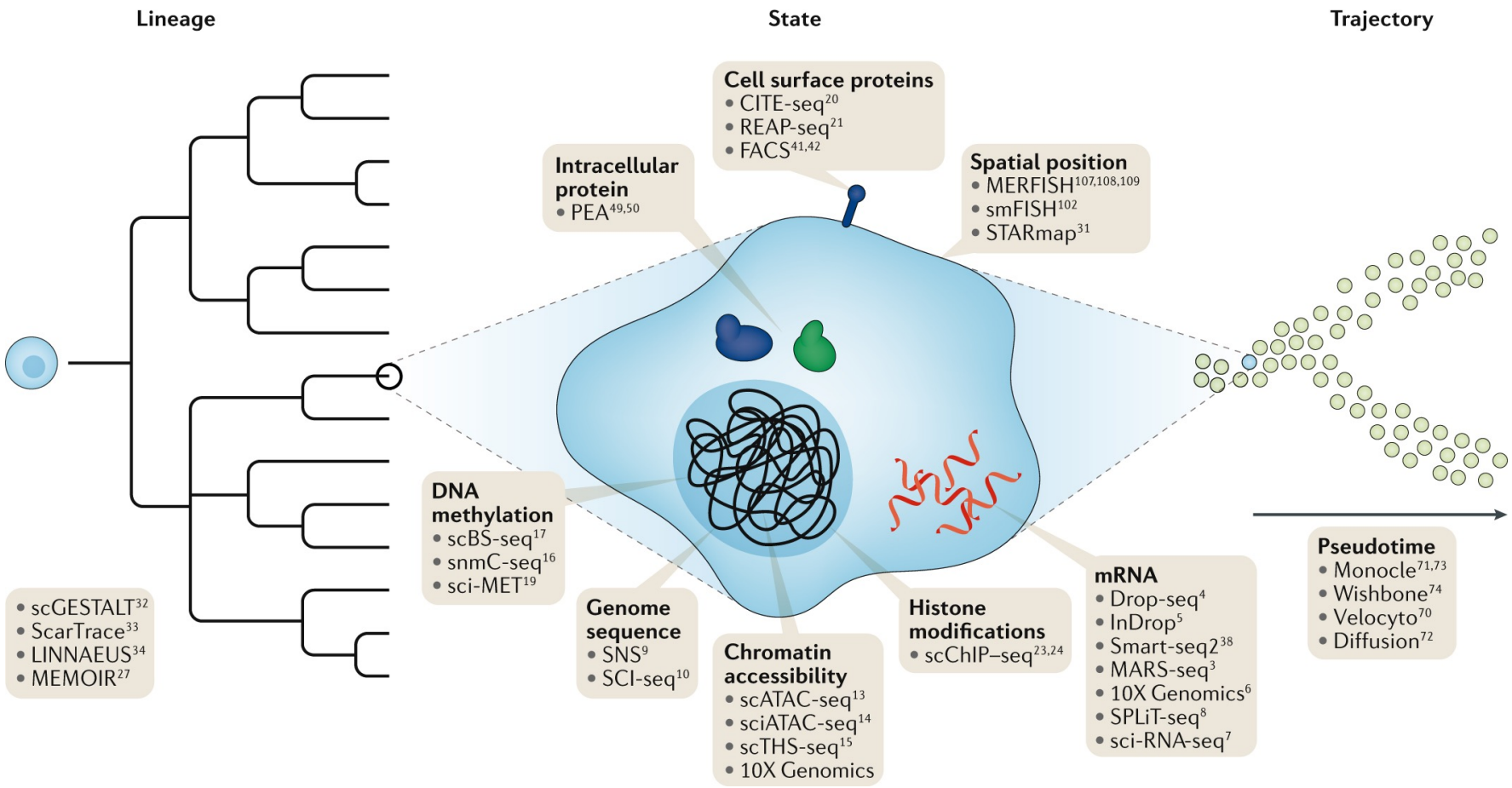
4,113 mouth-pipetted cells!

The Rapid Rise of Single Cell Biology



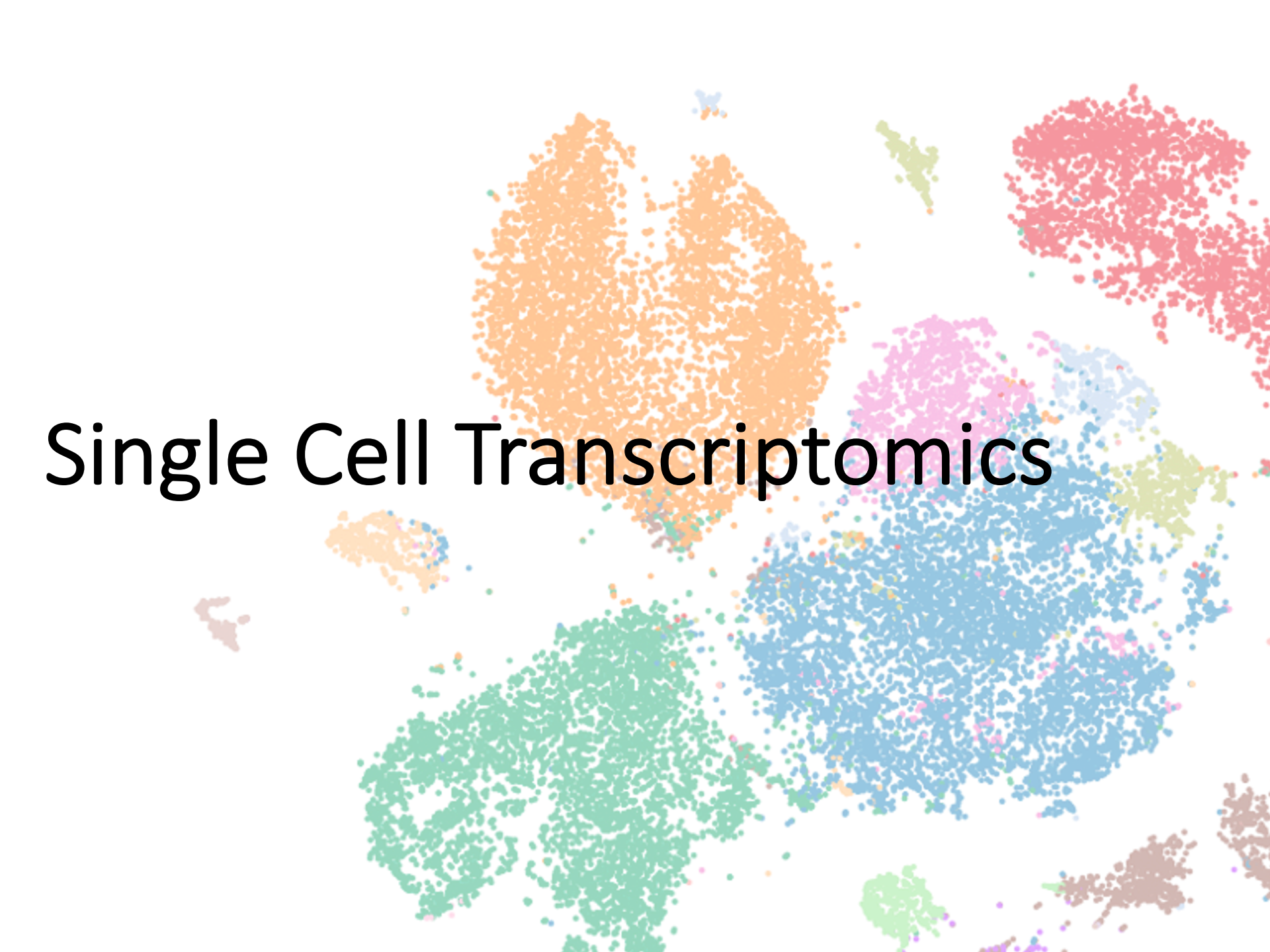
~10-fold increase in # of cells profiled every other year

Many Flavors of Single cell 'Omics

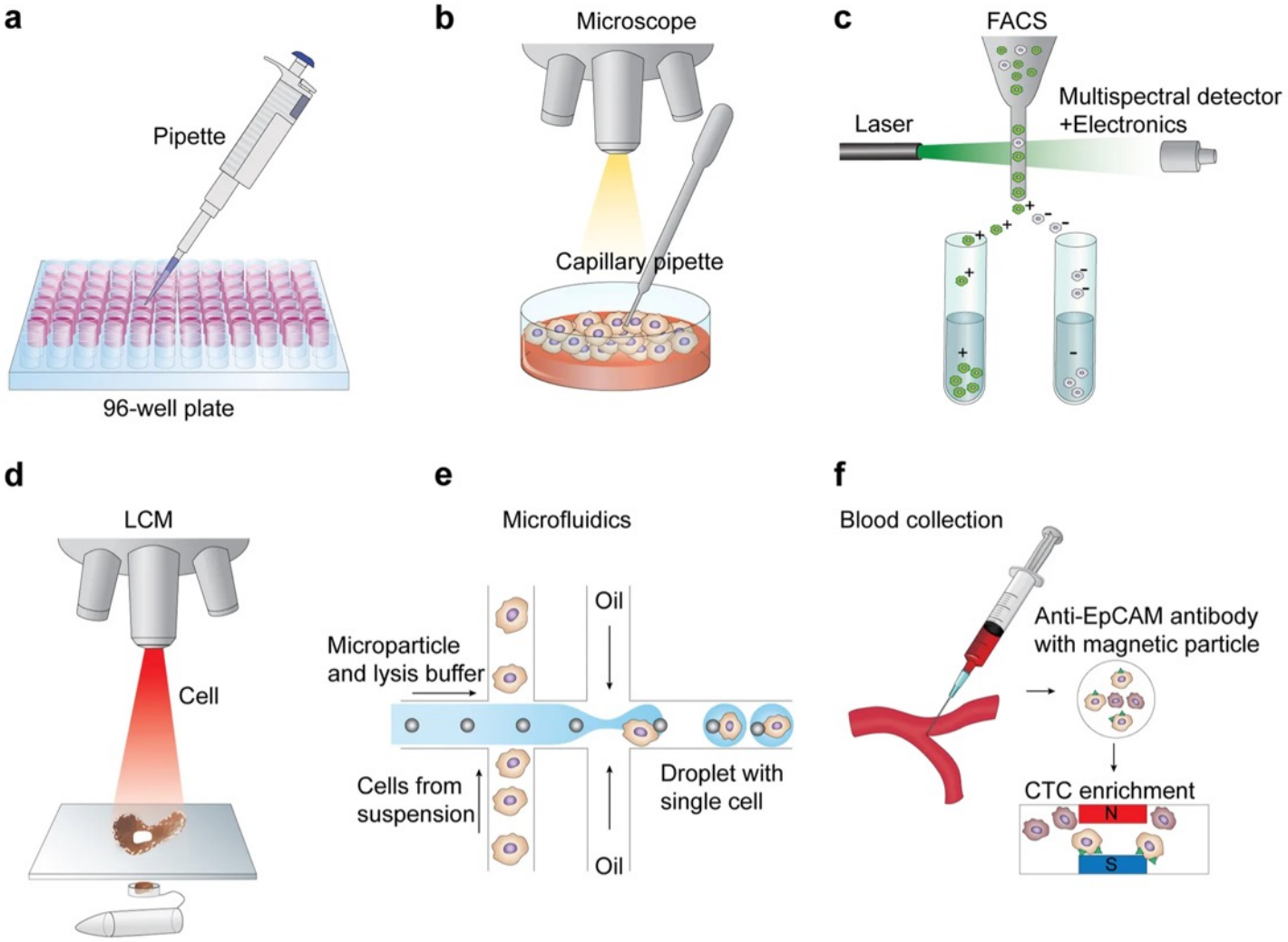


Stuart, T., Satija, R. Integrative single-cell analysis. *Nat Rev Genet* **20,** 257–272 (2019). https://doi.org/10.1038/s41576-019-0093-7

Single Cell Transcriptomics

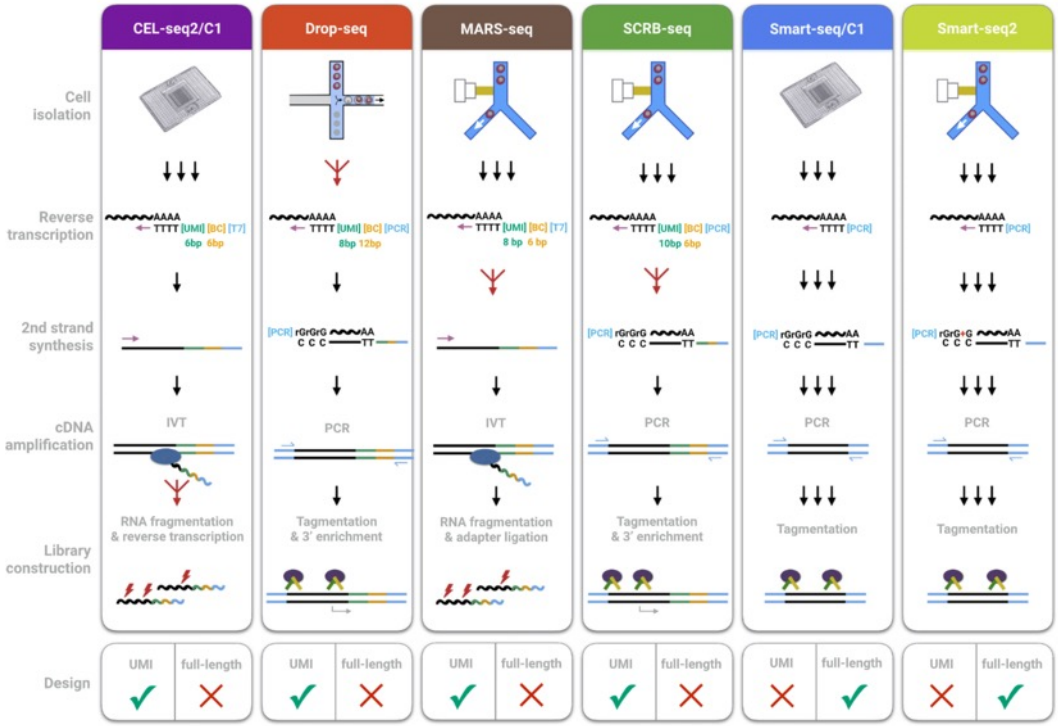


Step 1: Partitioning Cells



Hwang et al. *Ex Mol Med* **50**,(2018)

Step 2: Library Preparation



What question are you asking?

Simple Gene expression?

Strand-selective?

Alternative splicing / polyA / TSS?

Allele-specific expression?

Genotype heterogeneity (eg. in cancer)?

Depth vs Breadth?

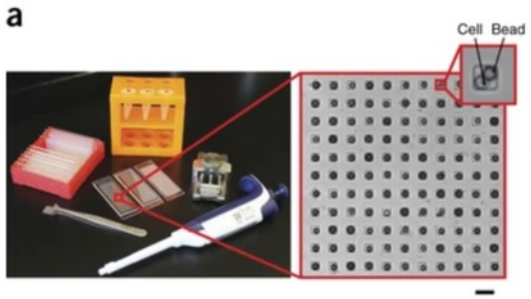
Ziegenhain et all. *Mol Cell* **65:**4, (2017), 631-643.e4

Most Common Platforms

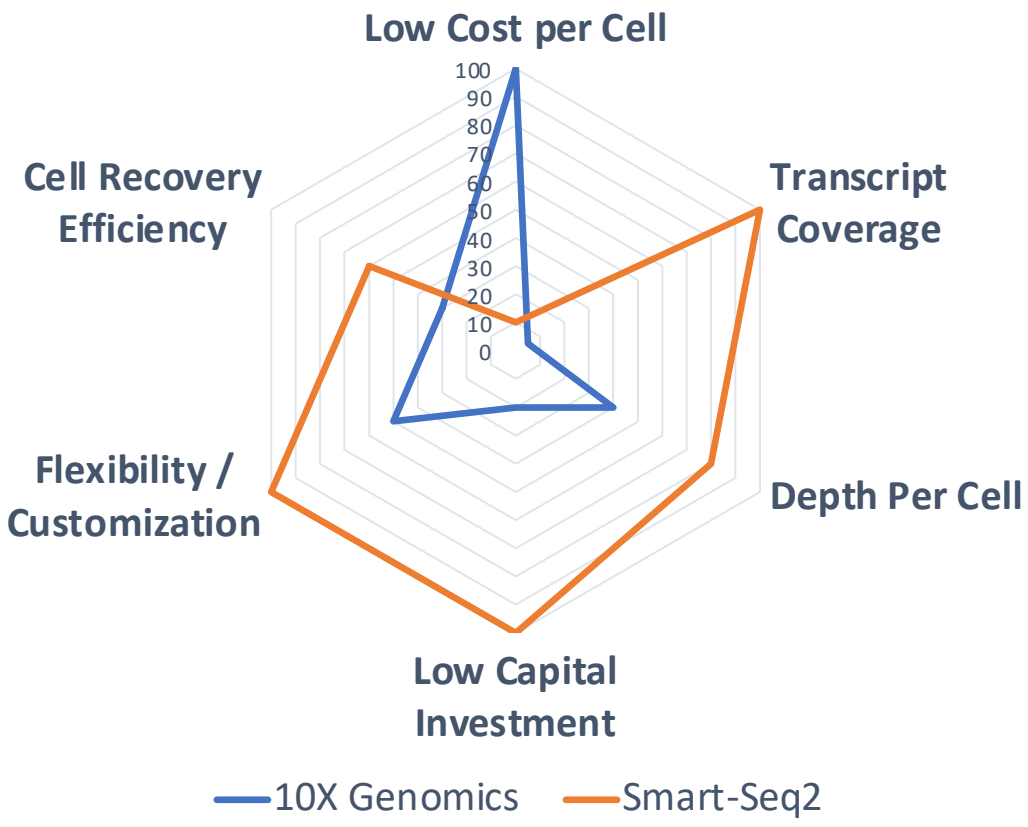
- Droplet / Bead
  - 10X Genomics Chromium
  - BD Rhapsody
  - Bio-Rad ddSeq

- Plate-based
  - SMART-Seq (v2, v3)
  - CEL-Seq2

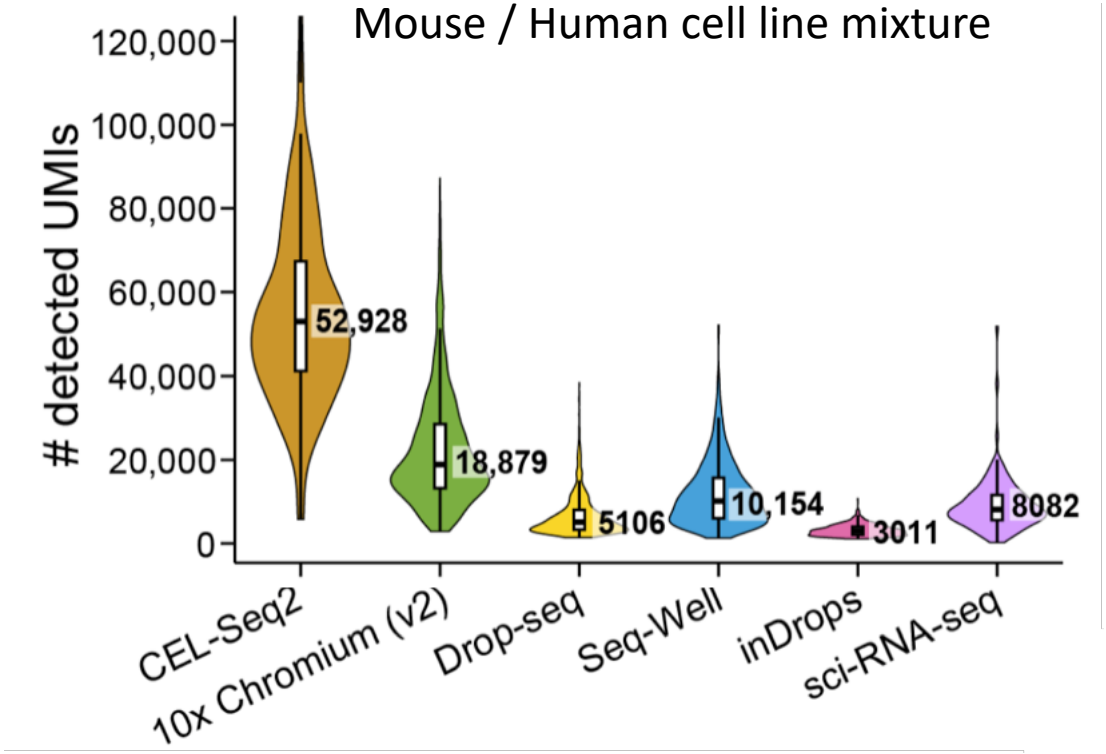
- Nanowell
  - Seq-Well



Which Method Should I Use?



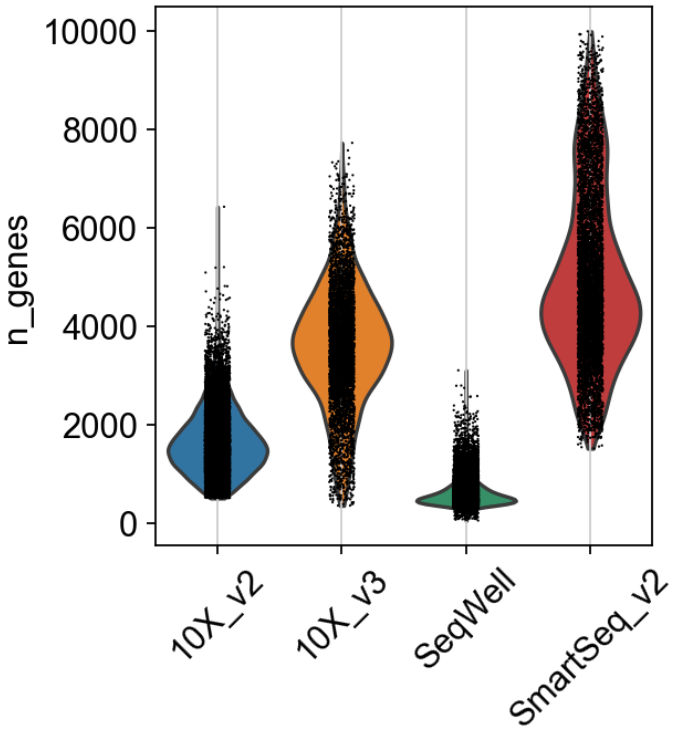
Systematic comparative analysis of single cell RNA-sequencing methods



Ding et al. bioRxiv (2019) https://doi.org/10.1101/632216

Mouse Fibroblasts:

Unique Genes Detected across technologies

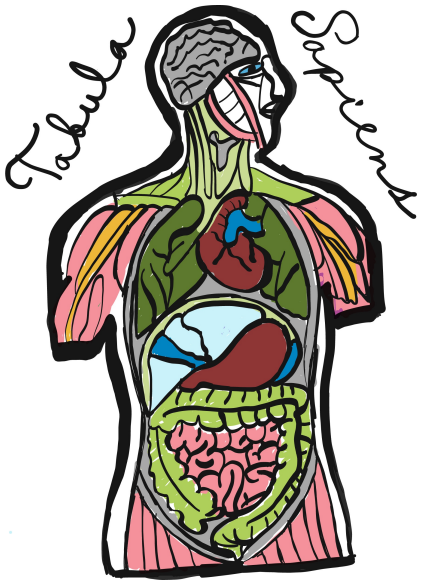


10X Genomics:
the *lingua franca* of the single-cell age

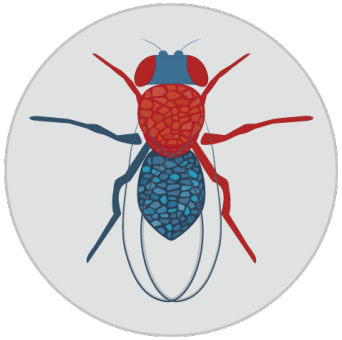
- Easy
- Robust
- Expensive.

COVID-19 Data Portal

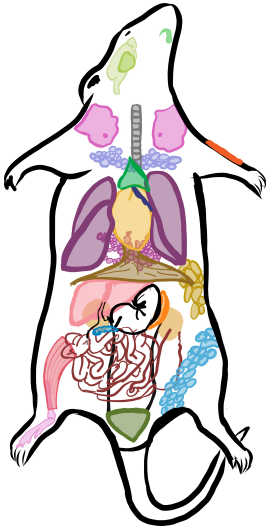
Allen Brain Map



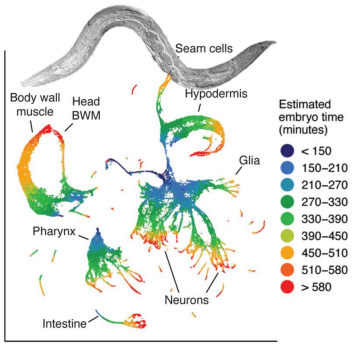
Fly Cell Atlas



Tabula Muris

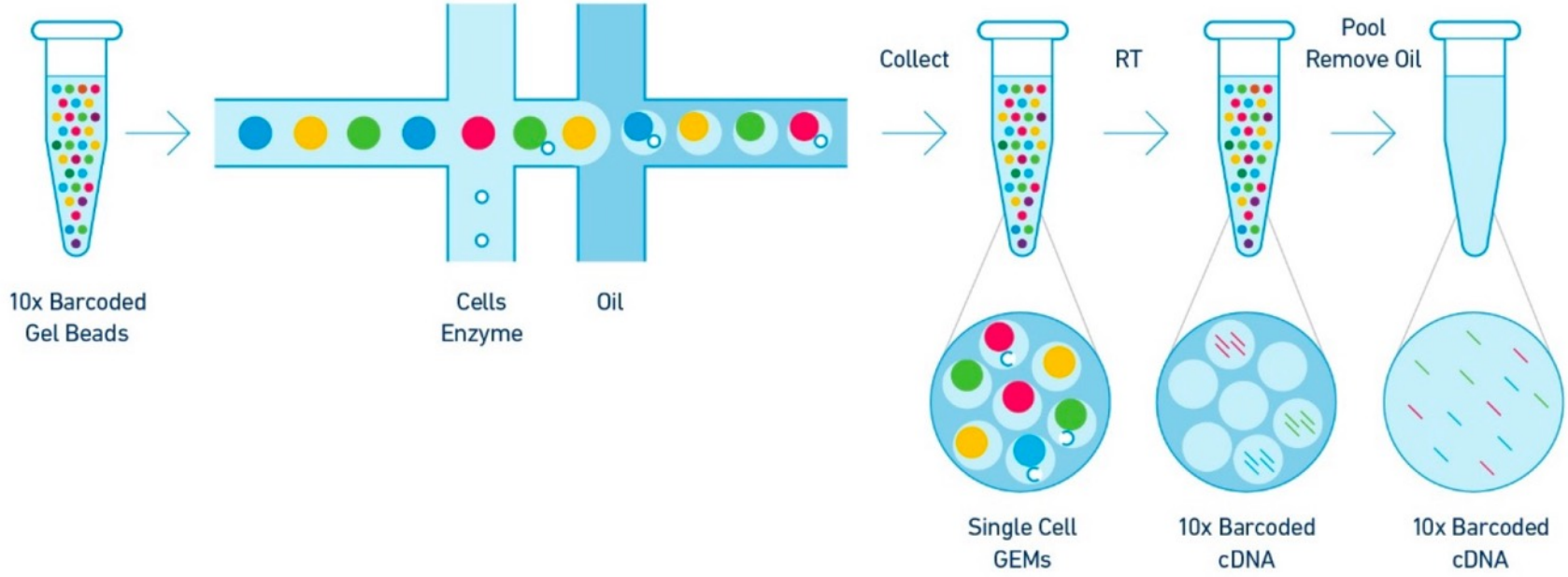


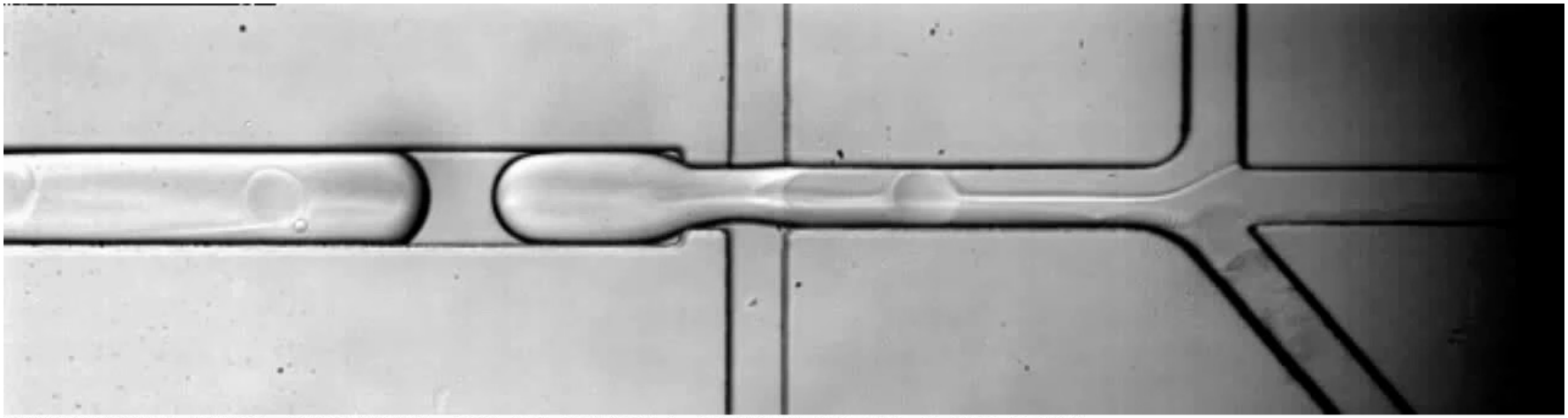
C elegans



Packer et al (2019) Science

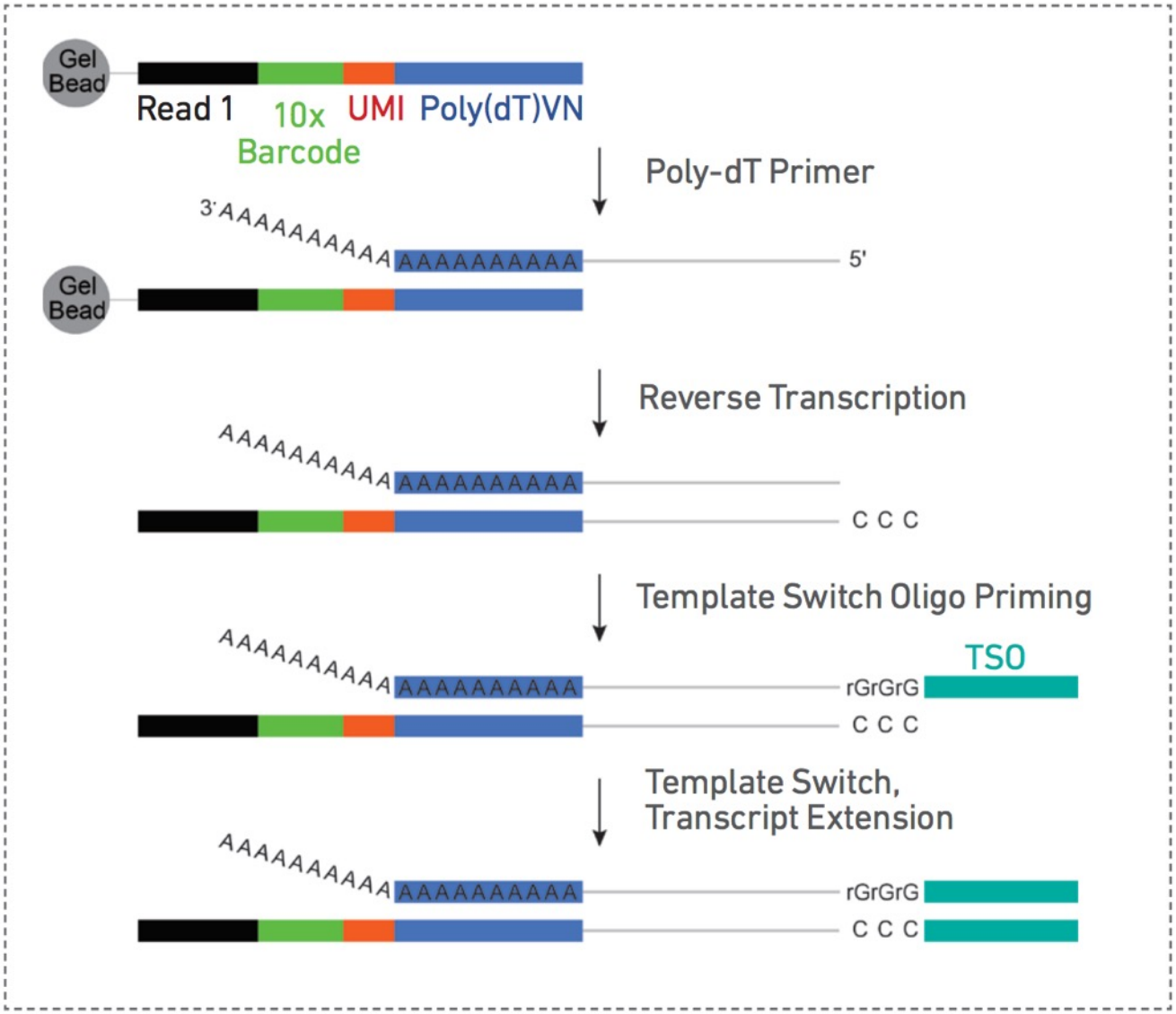
10X Genomics Workflow

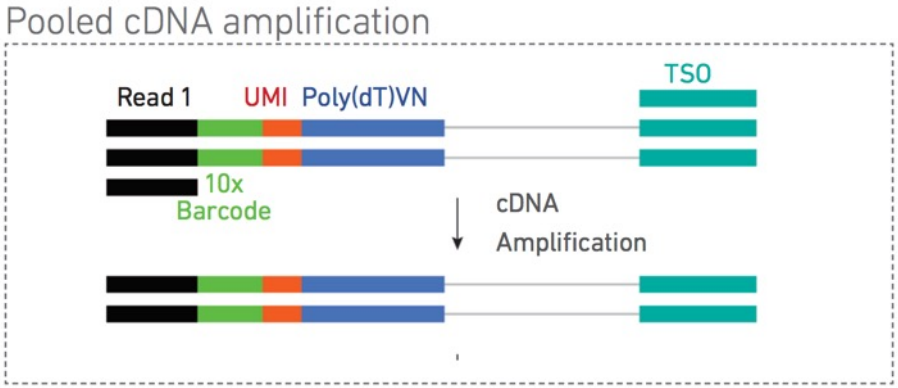


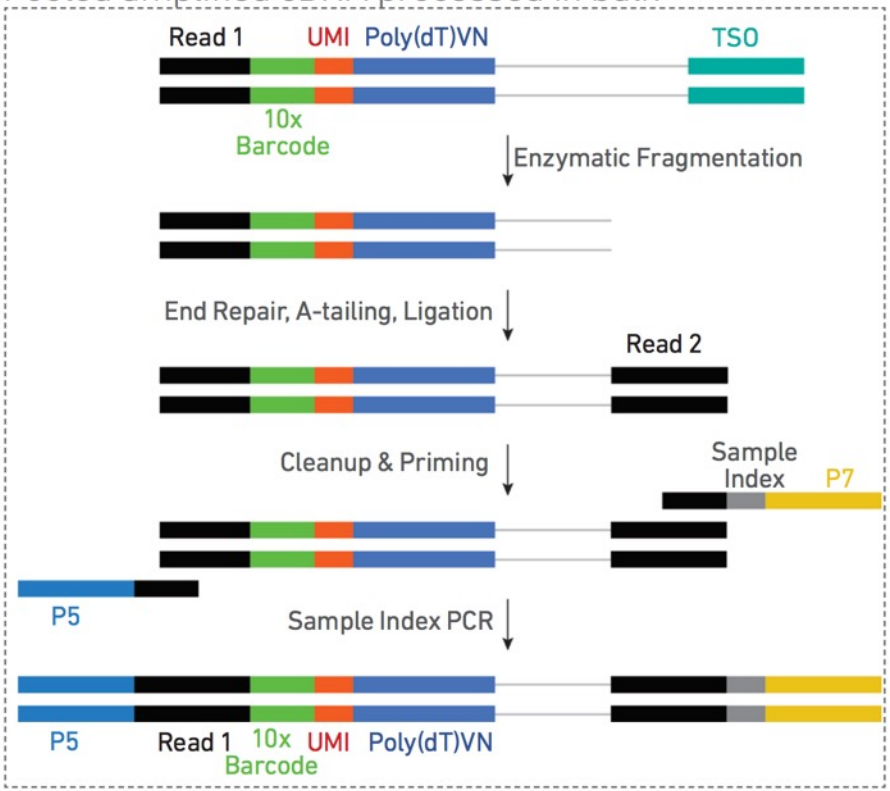


6/3/2014 9:30:12 PM -43738.7[ms] 000000523 HiSpec 1 [00-11-1c-f1-73-f3] Fastec 1280x336(Q) 400fps 100µs V1.4.3 (Build: 2419)

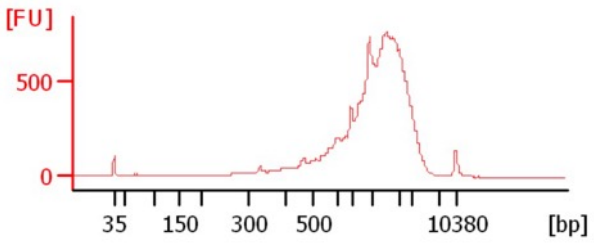
Single Cell 3’ Chemistry Overvivew



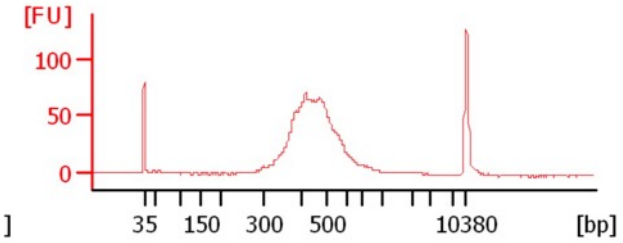




Bioanalyzer



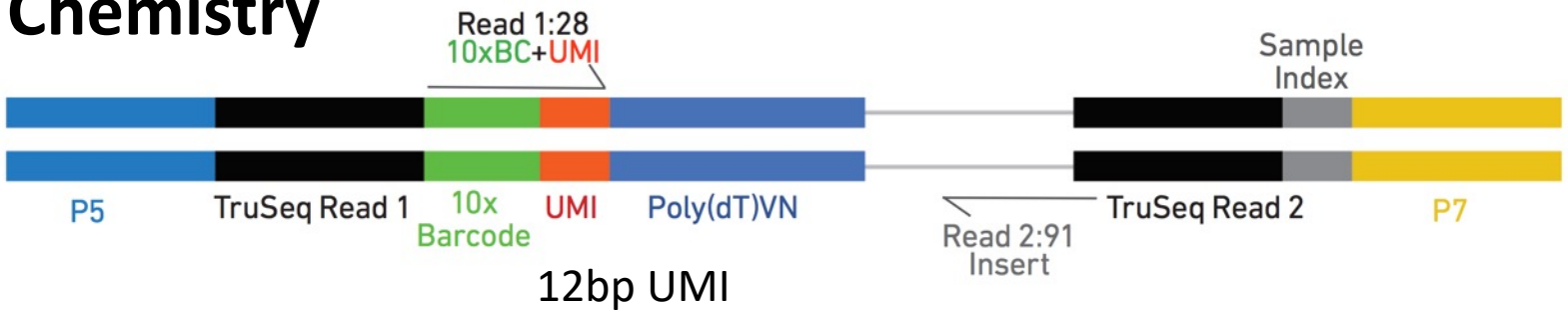
Amplified cDNA



Final Library

Anatomy of a 10X 3’-Single Cell Amplicon

V3 Chemistry



12bp UMI

Unique Molecular Identifier (UMI)

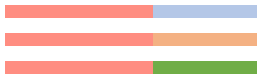
Random ~8-10bp sequence incorporated during oligo synthesis

Cell barcode UMI

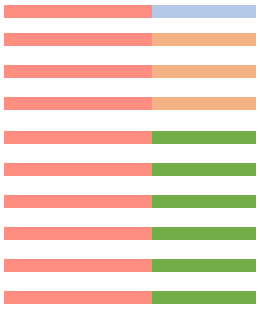
CCCCCCCCXXXXXXTTTTTTTTTTVN
AAAAAAAAAABN---IFNgamma----

CCCCCCCCXXXXXXTTTTTTTTTTVN
AAAAAAAAAABN---IFNgamma----

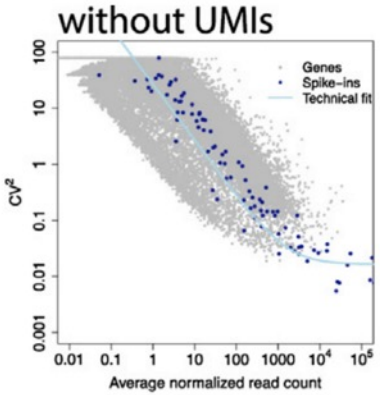
CCCCCCCCXXXXXXTTTTTTTTTTVN
AAAAAAAAAABN---IFNgamma----

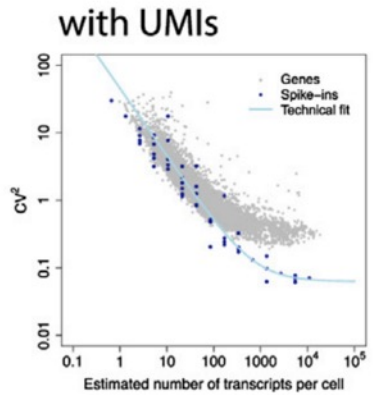


Before PCR



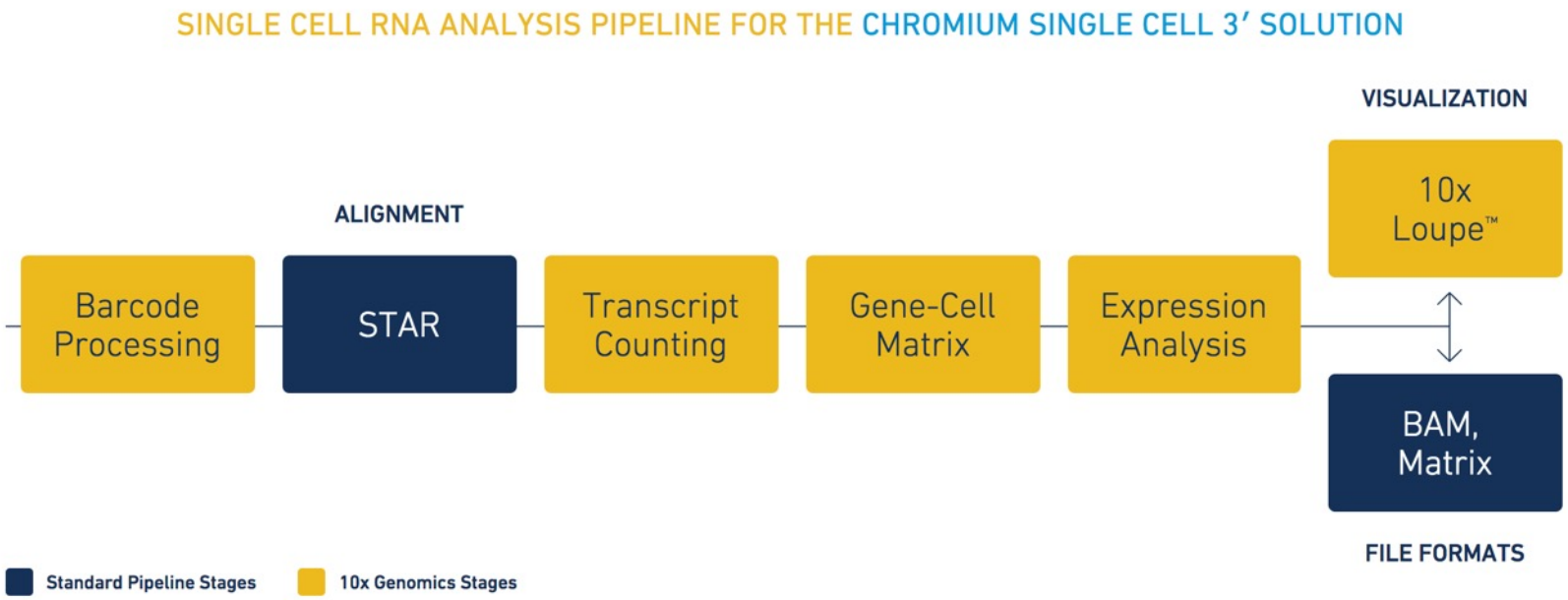
After PCR





Mapping and Transcript Quantification

Cellranger Count pipeline: [10X Genomics support page]



Digital Gene Expression, Not Coverage

“Deep” Single Cell Libraries

Well-based, eg. SmartSeq

Fluidigm C1

Droplet – Based DGE libraries

Drop-Seq

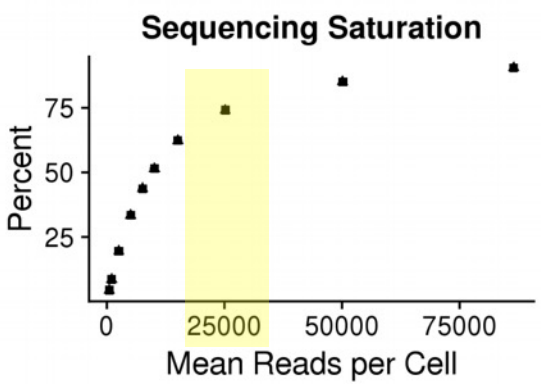
10X Genomics

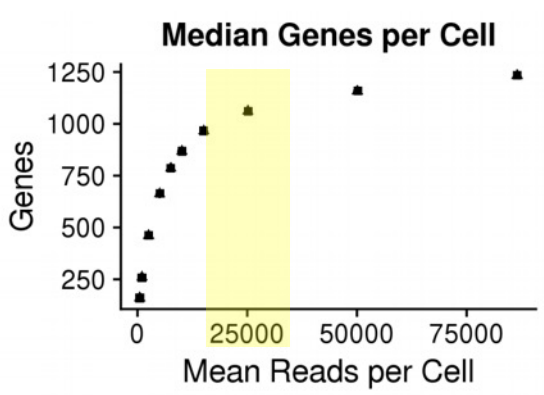
Seq-Well

How Deeply Should I Sequence?

Human PBMC

Subsampled Library

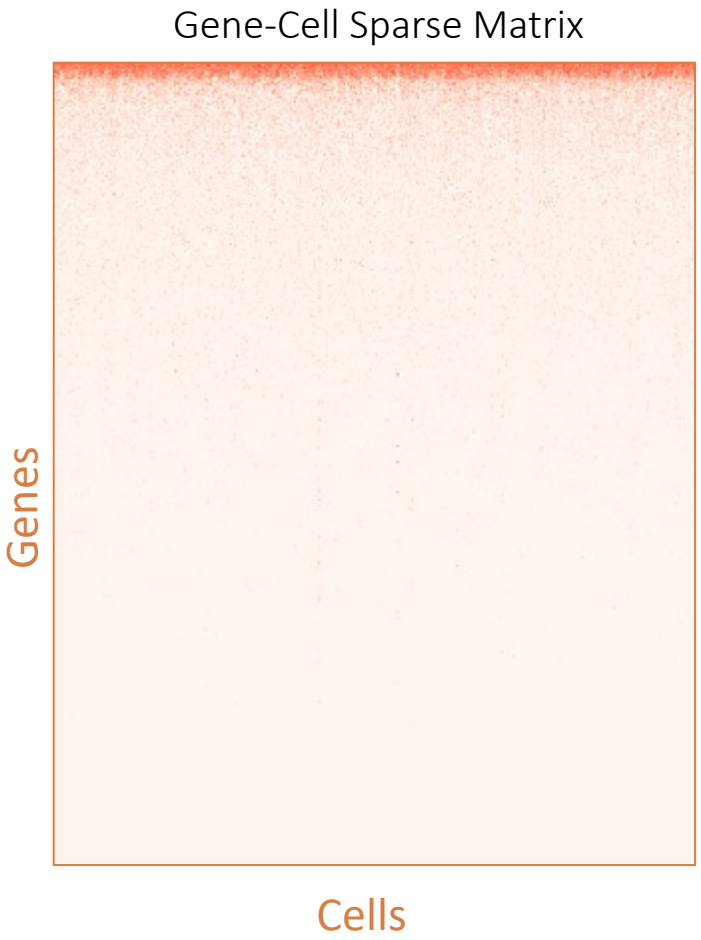




Median Reads / Cell

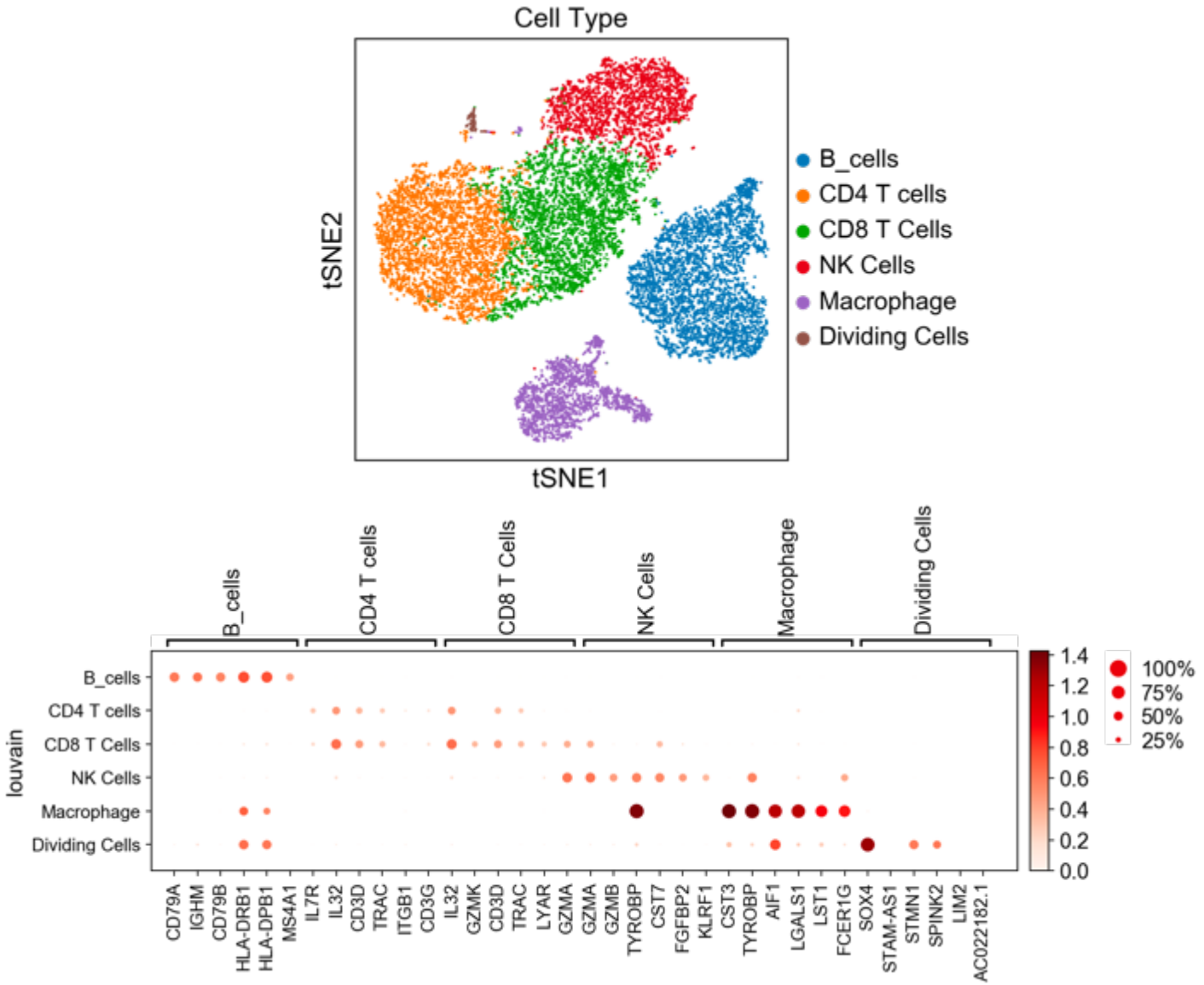
500 | 5,000 | 25,000 | 86,503

Sparse sampling of gene expression

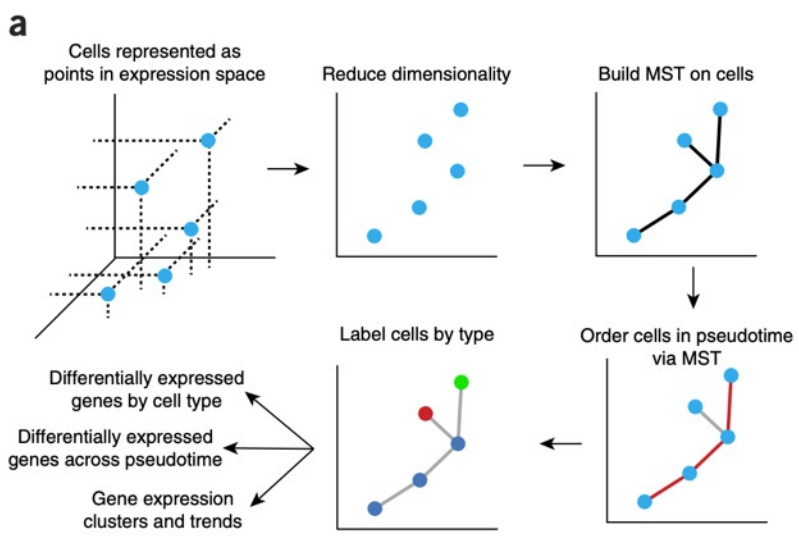


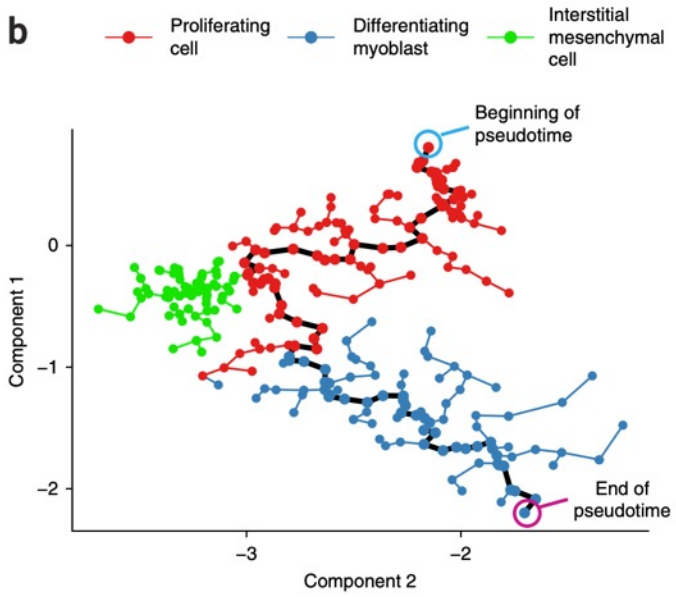
Top	Gene Expression	US Wealth
1%	15%	35%
10%	55%	73%
20%	73%	86%

Basic output of scRNAseq pipeline

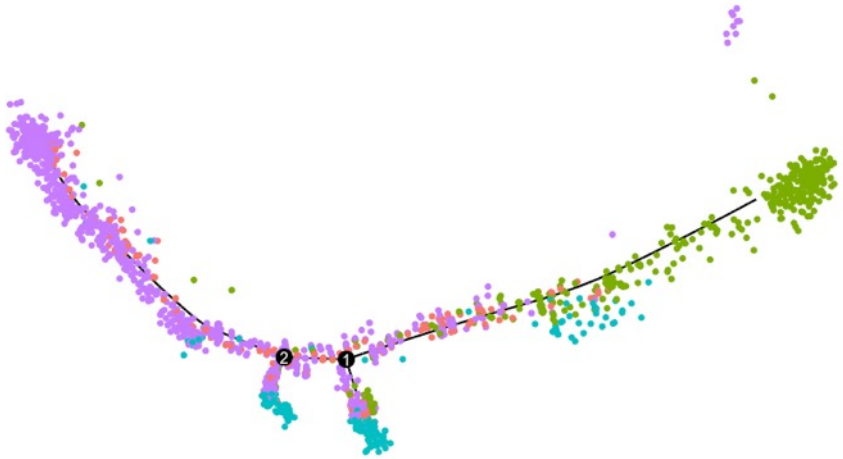


Pseudotime analysis





Monocle

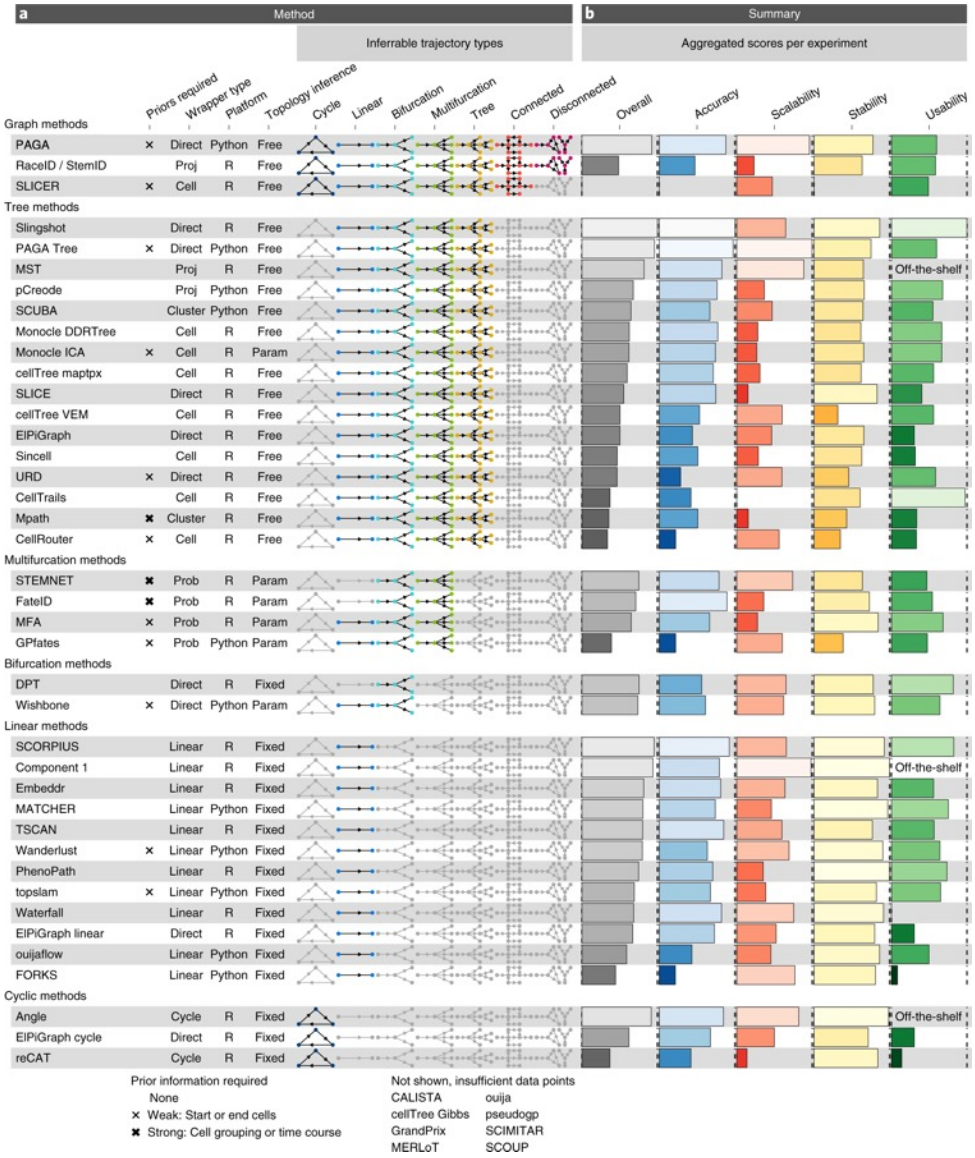


https://cole-trapnell-lab.github.io/monocle3/docs/installation/

Pseudotime

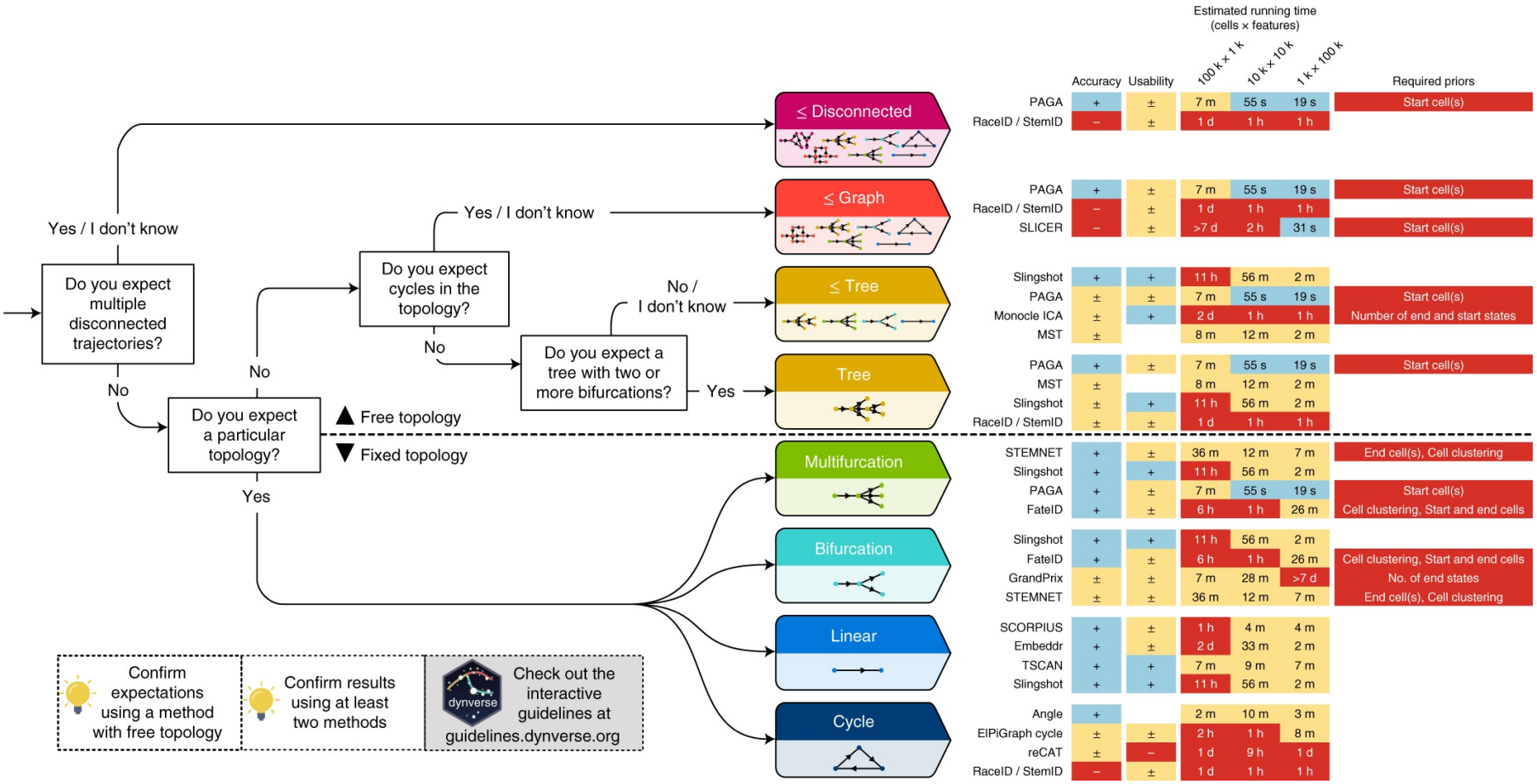
Dozens of methods developed

Vary in terms of feature selection, dimensionality reduction, tree construction, etc



Saelens, W., Cannoodt, R., Todorov, H. et al. A comparison of single-cell trajectory inference methods. Nat Biotechnol 37, 547–554 (2019). https://doi.org/10.1038/s41587-019-0071-9

Pseudotime – which method to use?



Saelens, W., Cannoodt, R., Todorov, H. et al. A comparison of single-cell trajectory inference methods. Nat Biotechnol 37, 547–554 (2019). https://doi.org/10.1038/s41587-019-0071-9

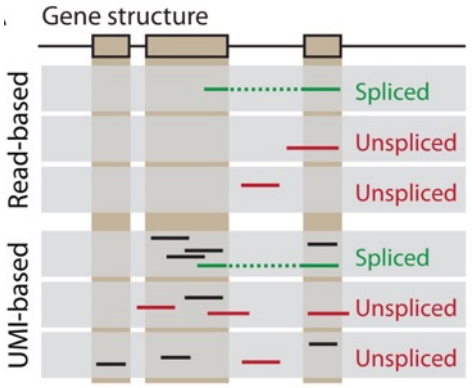
RNA Velocity

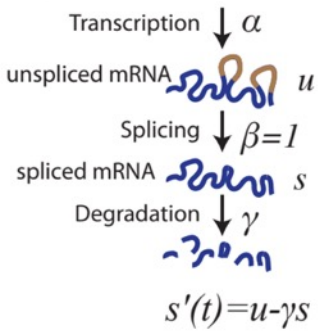
Estimates rates of change in mRNA levels by modeling nascent RNA synthesis

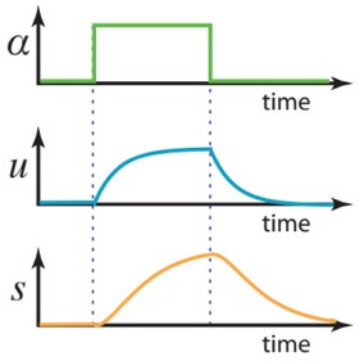
Quantifies spliced / unspliced

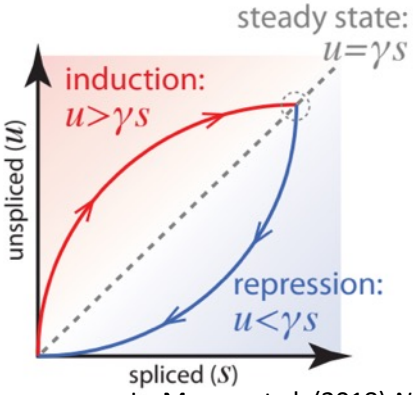
Models dynamics

CAVEATS: Gene annotations
Cryptic exons
unannotated intronic genes
repetitive elements





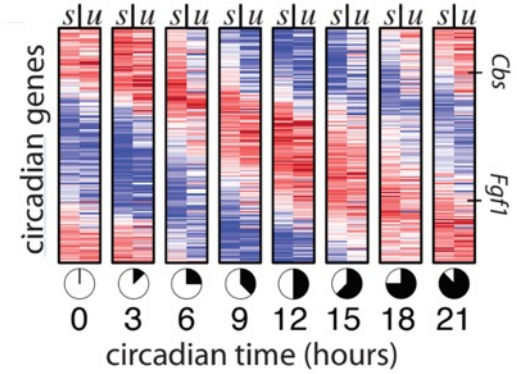


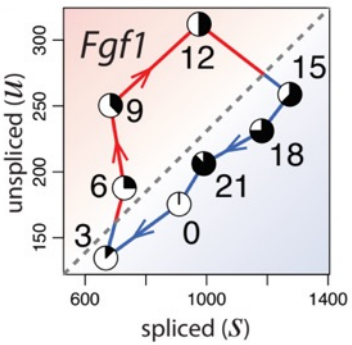


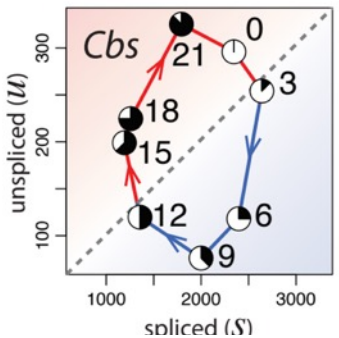
Le Manno et al. (2018) *Nature*

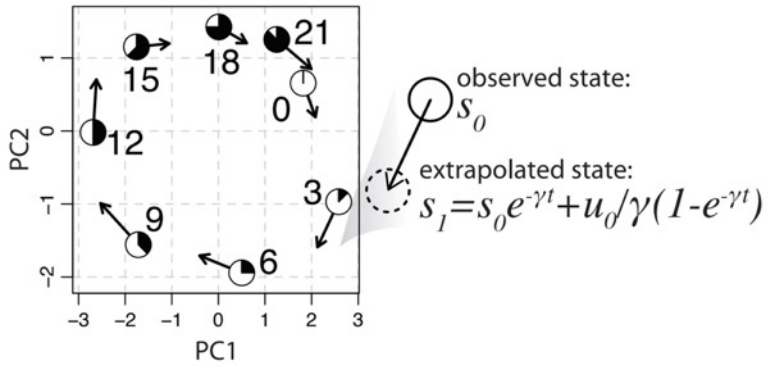
RNA Velocity

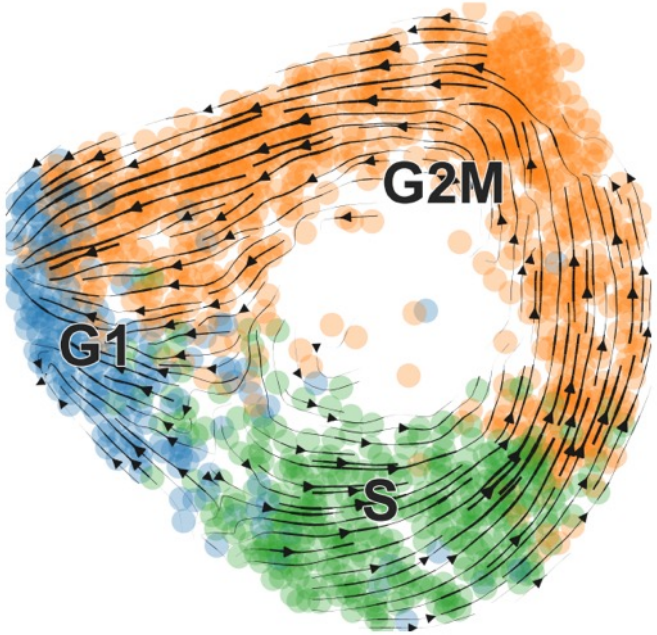
Bulk RNAseq from mouse circadian rhythm data



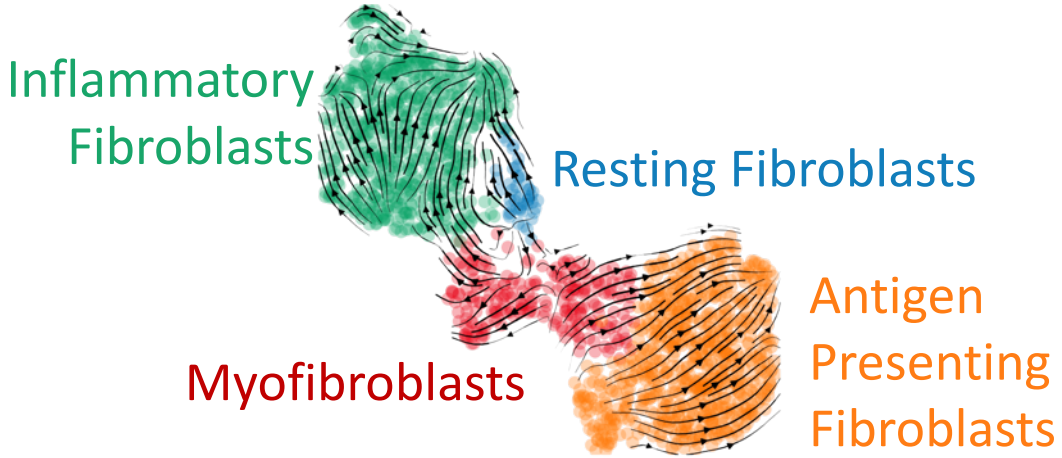


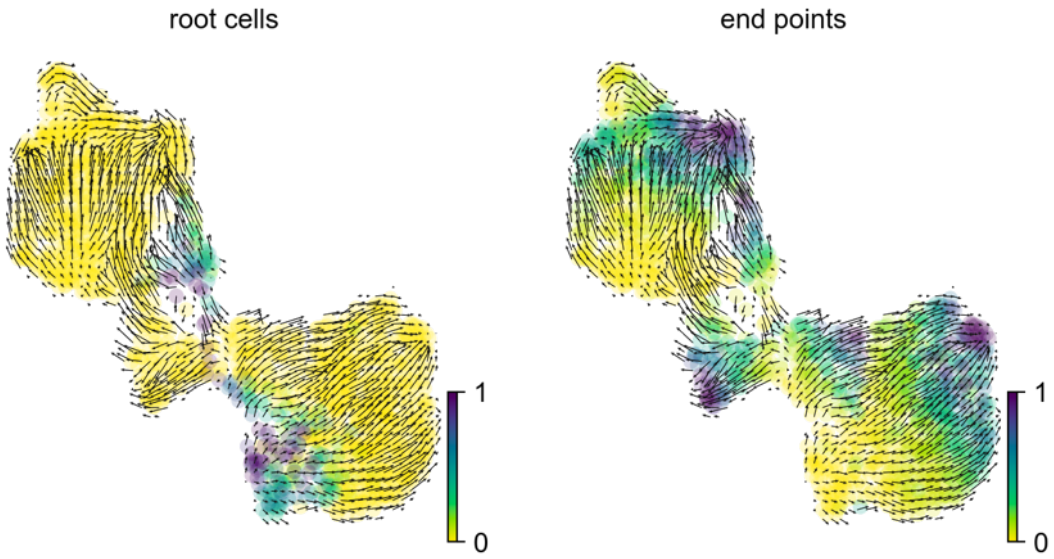


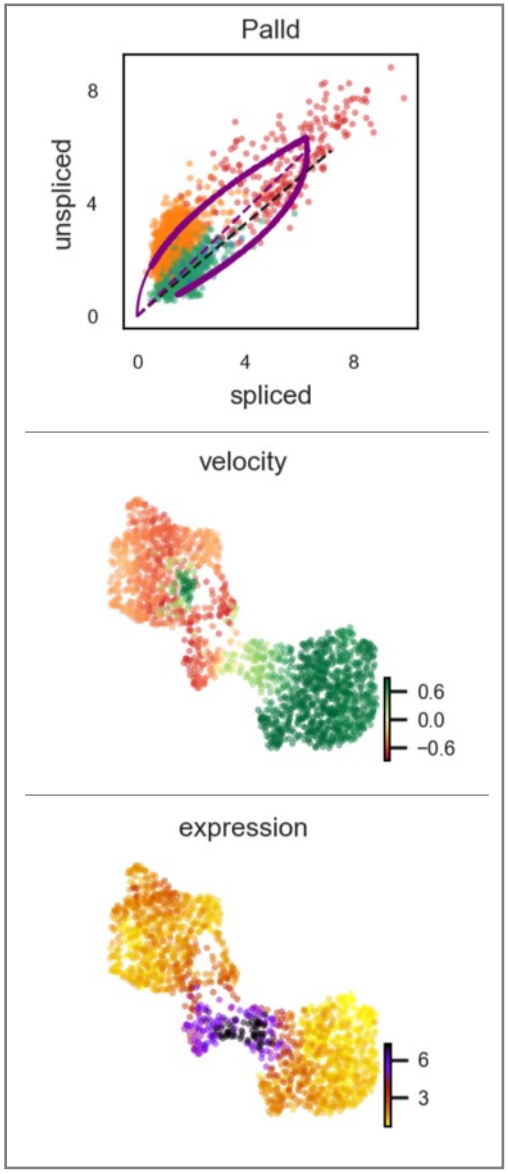




Inferring Differentiation Trajectories from RNA Velocity







SCENIC

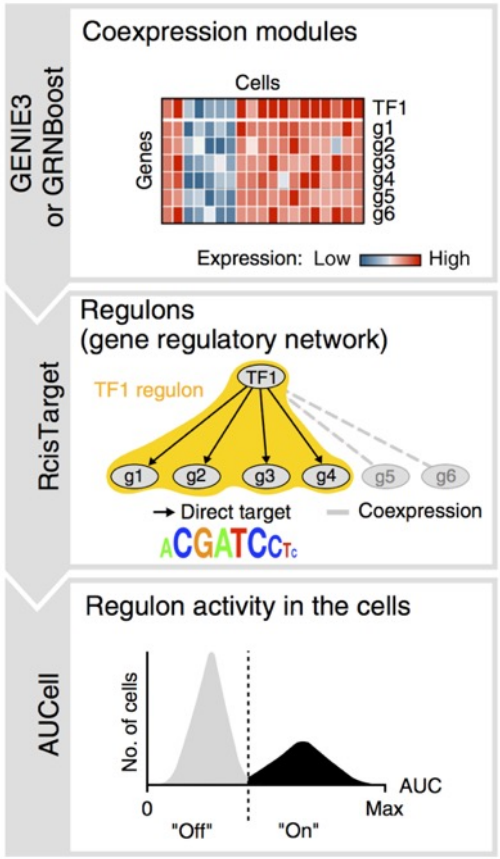
single-cell regulatory network inference and clustering

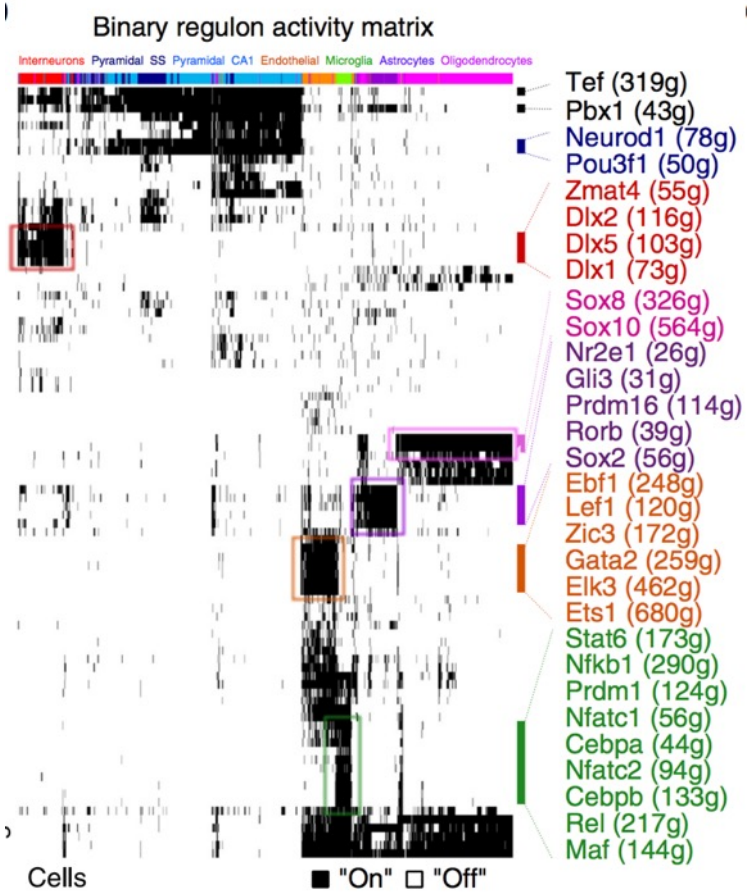
Transcription Factor Activity Inference

Gene Co-expression network

Motif search

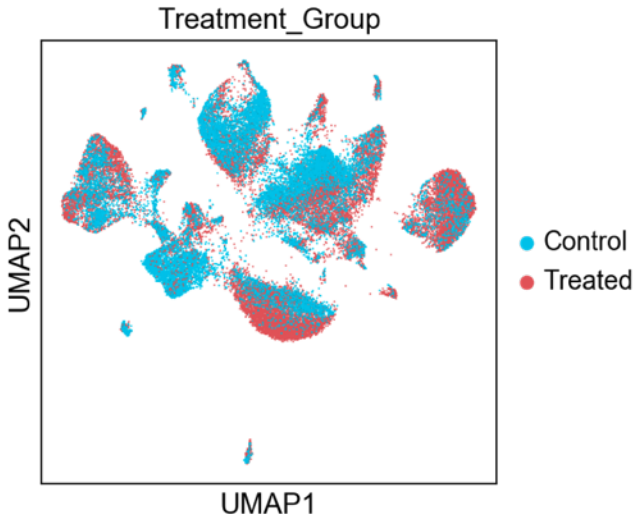
Regulon activity

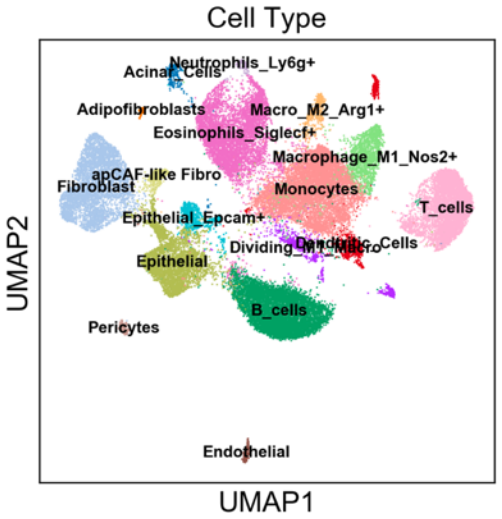


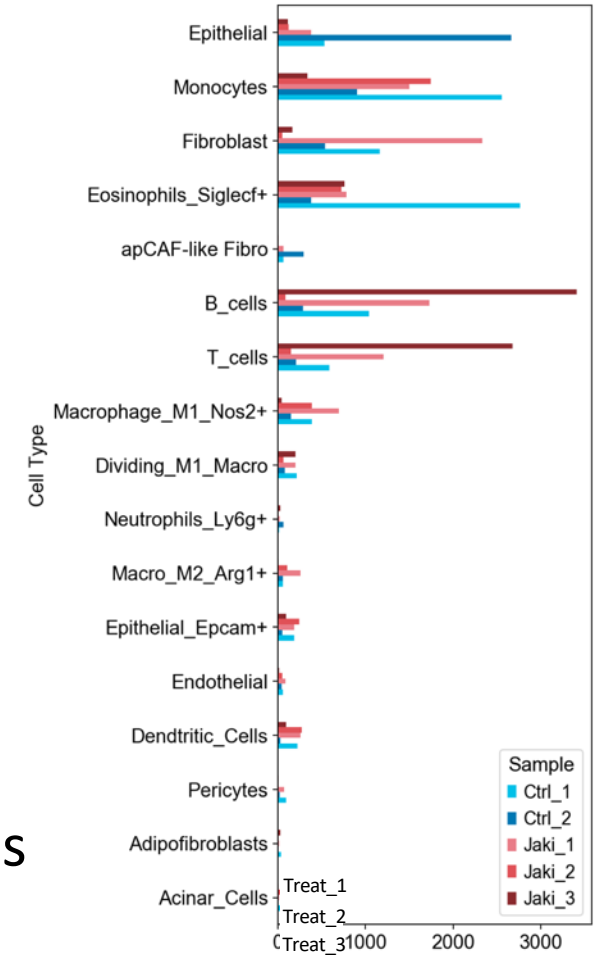


Aibar et al. (2017) *Nature Methods.* SCENIC: single-cell regulatory network inference and clustering

scRNAseq is a poor cytometry tool

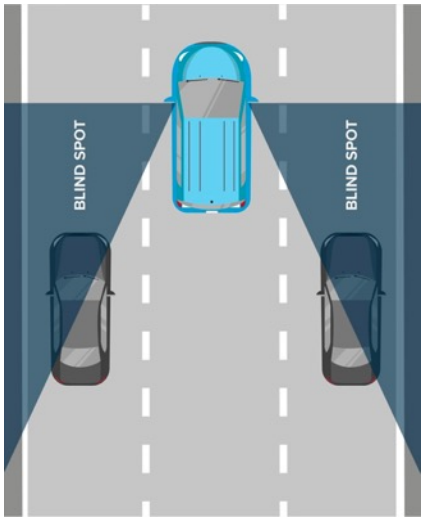






- Unreliable – highly sensitive to conditions
- Expensive
- Low throughput

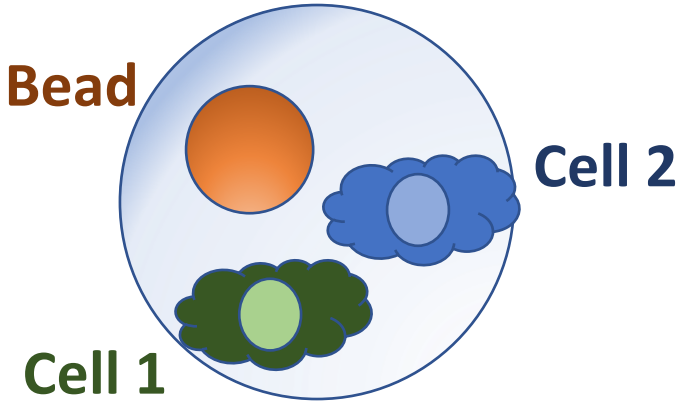
Blind Spots



- Some cell types might be missed
  - Low mRNA count – filtered from matrix
    - Early 10X Genomics Software (v2)
    - Defaulted to exclude lots of lymphocytes
  - Hard to dissociate from tissue
    - Fibroblasts
  - Cells might die quickly during prep
    - Stem cells
  - Fragile: (Acinar cells, Plasma cells)
  - High RNAse / protease content (Acinar, Neutrophils)
    - Peripheral blood neutrophils especially!!!
  - Doublets / Multiplets

Doublets

- Proportional to concentration of cell suspension



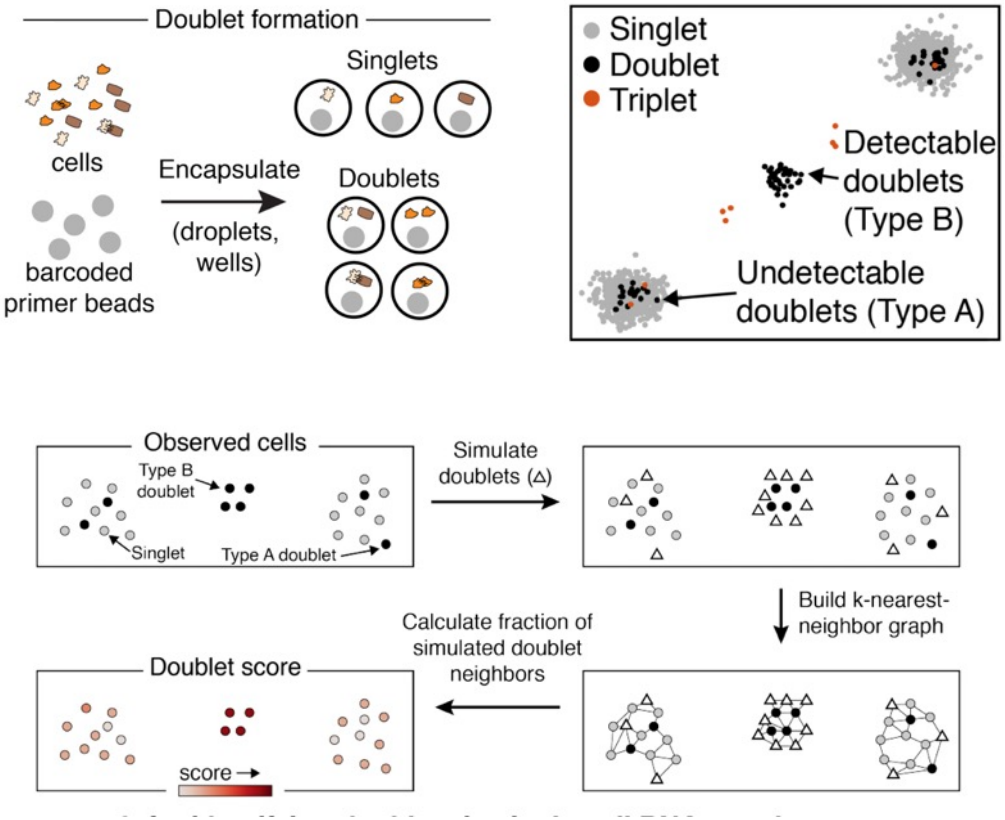
•DoubletFinder - [R] - Doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. BioRxiv

•DoubletDecon - [R] - Cell-State Aware Removal of Single-Cell RNA-Seq Doublets. [BioRxiv](DoubletDecon: Cell-State Aware Removal of Single-Cell RNA-Seq Doublets)

•DoubletDetection - [R, Python] - A Python3 package to detect doublets (technical errors) in single-cell RNA-seq count matrices. An R implementation is in development.

•Scrublet - [Python] - Computational identification of cell doublets in single-cell transcriptomic data. BioRxiv

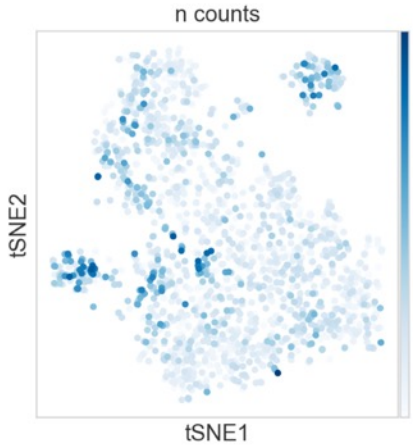
Scrublet



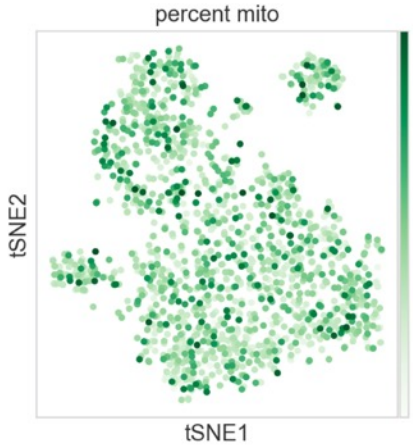
Wolock et al. (2018) bioRxiv

Sources of Measurement Noise

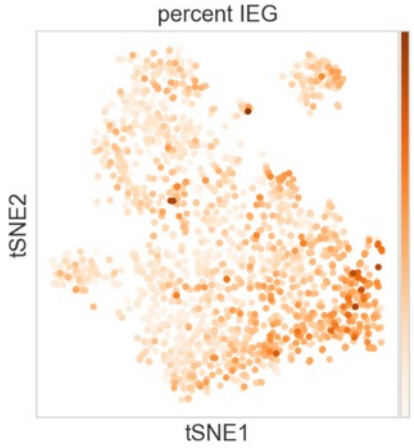
Library Depth



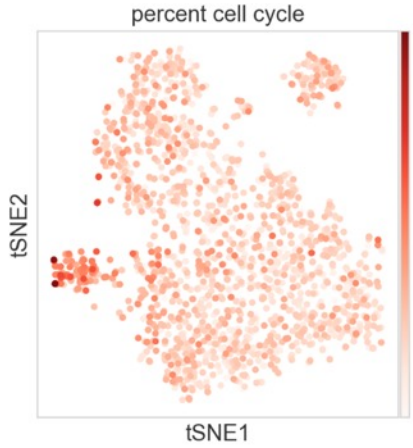
Cell Viability



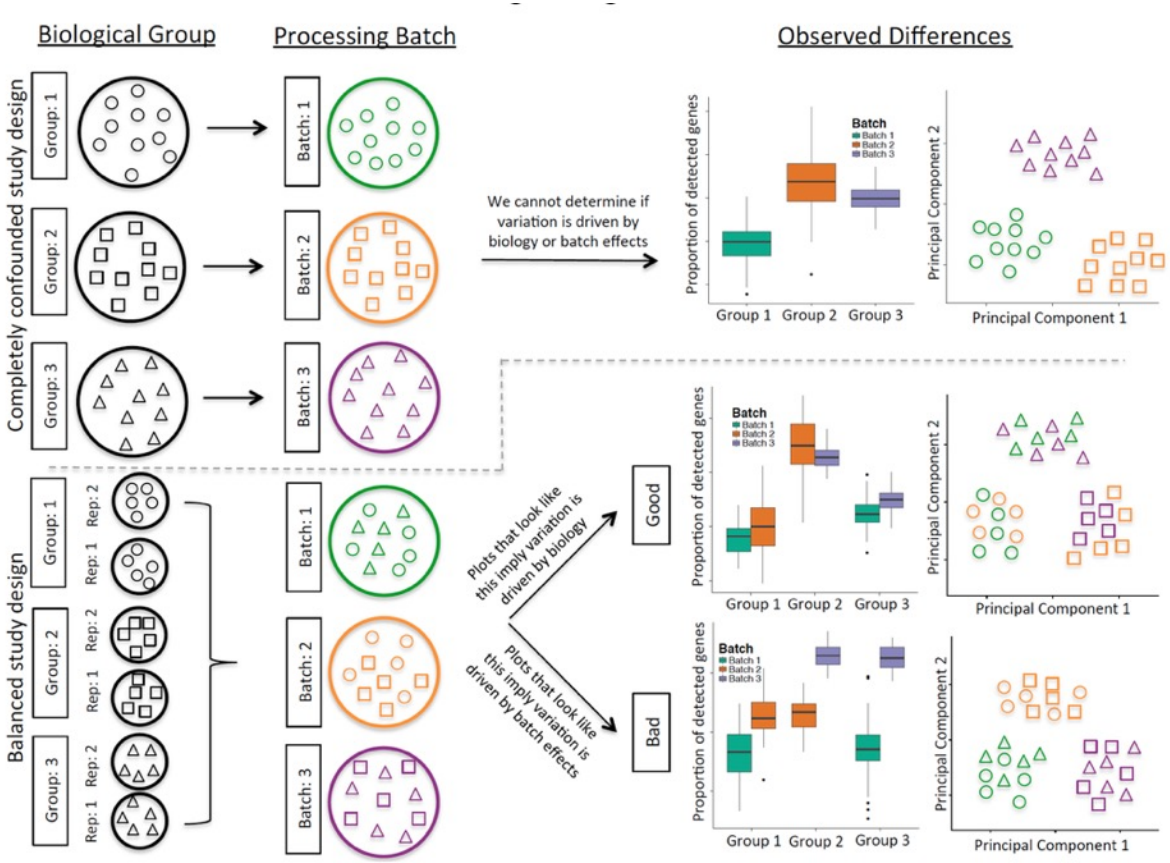
Dissociation Artifacts



Cycling Cells

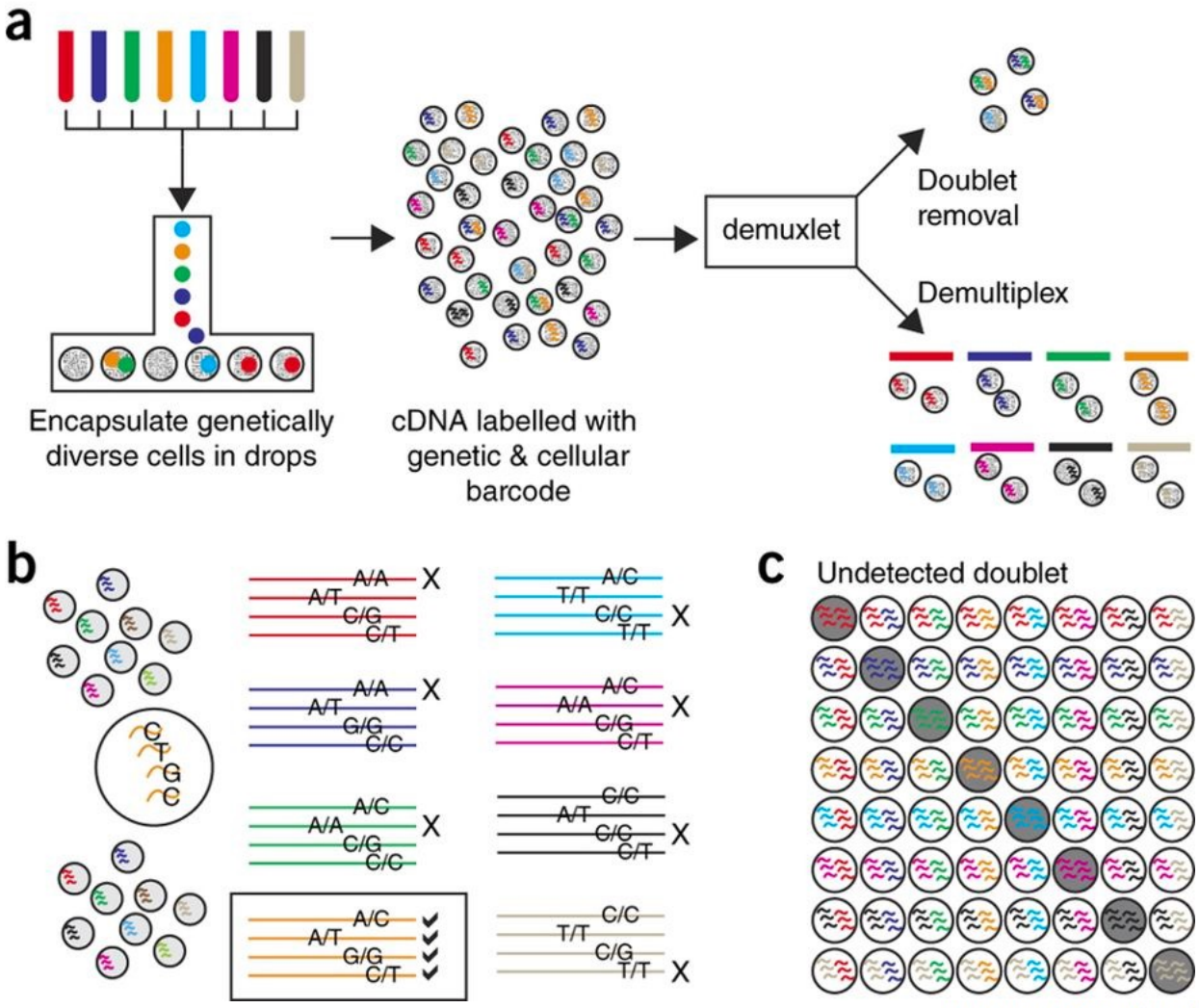


Batch effects and study design



Multiplexing Using Natural Genetic Variation

Demuxlet



Jimmie Ye lab

Nature Biotechnology **36**, 89–94 (2018)

Sex – matched studies are helpful!

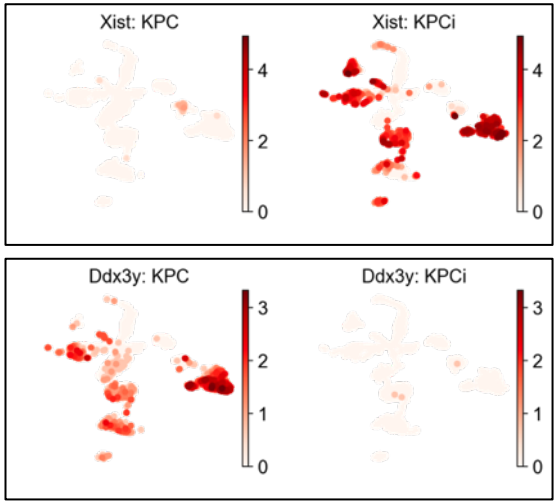
Major confounder: Male / Female

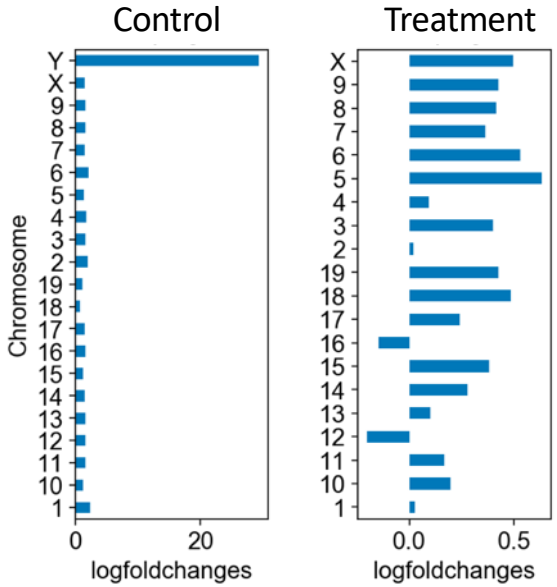
Treatment: Female

Control: Male

Consequence:

Unsupervised differential gene expression calling will be dominated by sex-specific expression. No way of separating this variable from the treatment variable





Batch Correction

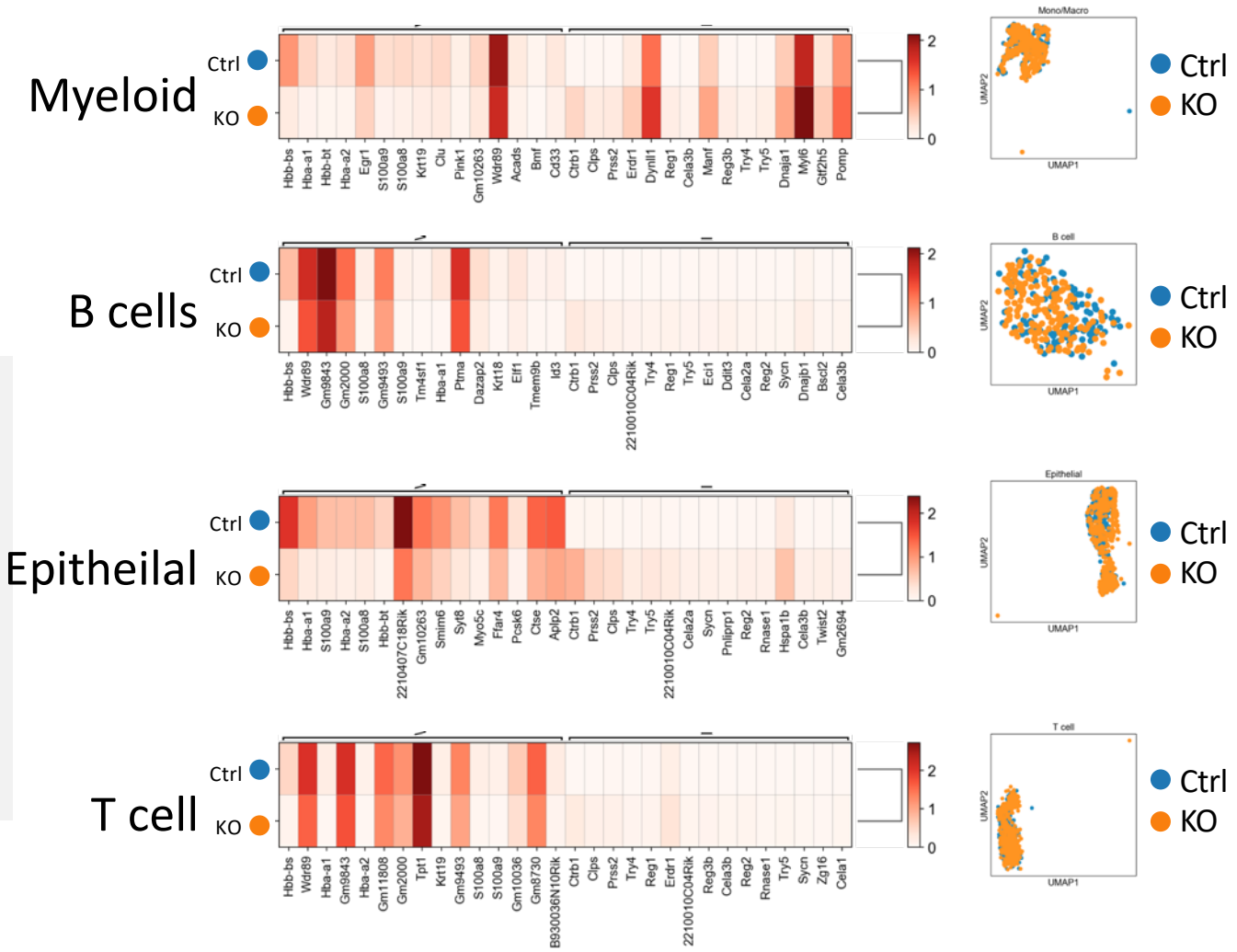
Confounded Study Example:

WT and KO mice

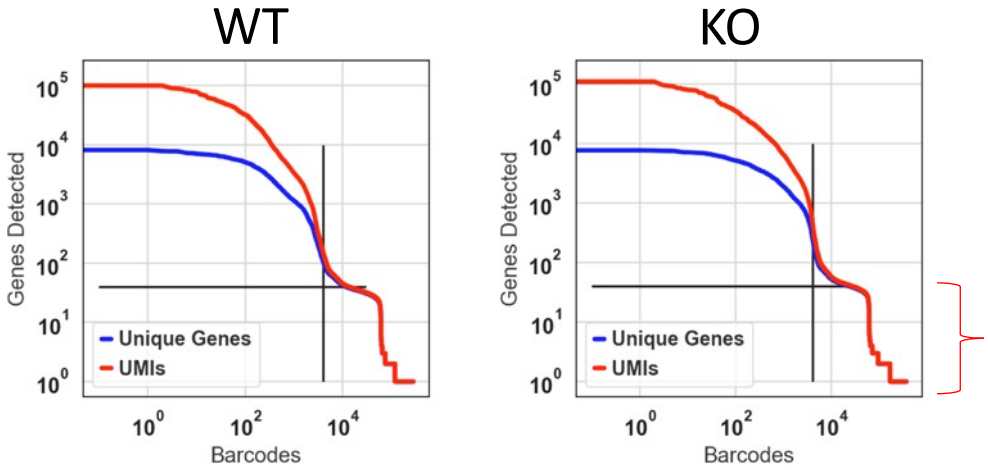
Prepared on same day
Same colony
Same set of hands

Diffex dominated by same genes within every cluster

! major batch effect issues

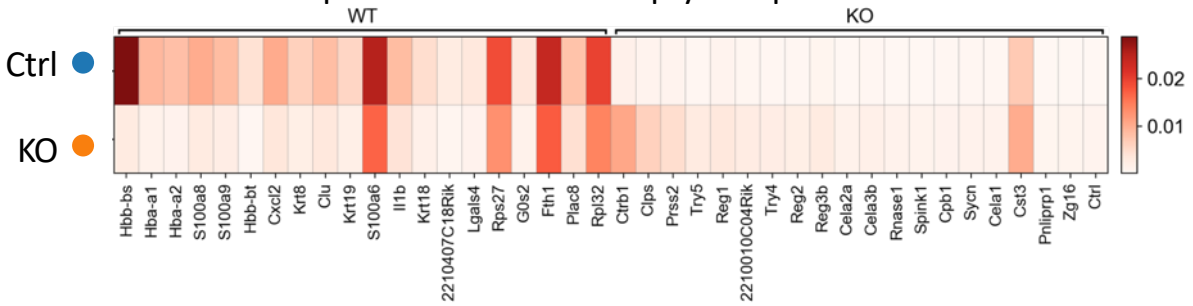


Controlling for batch effects



Ambient RNA
in droplets

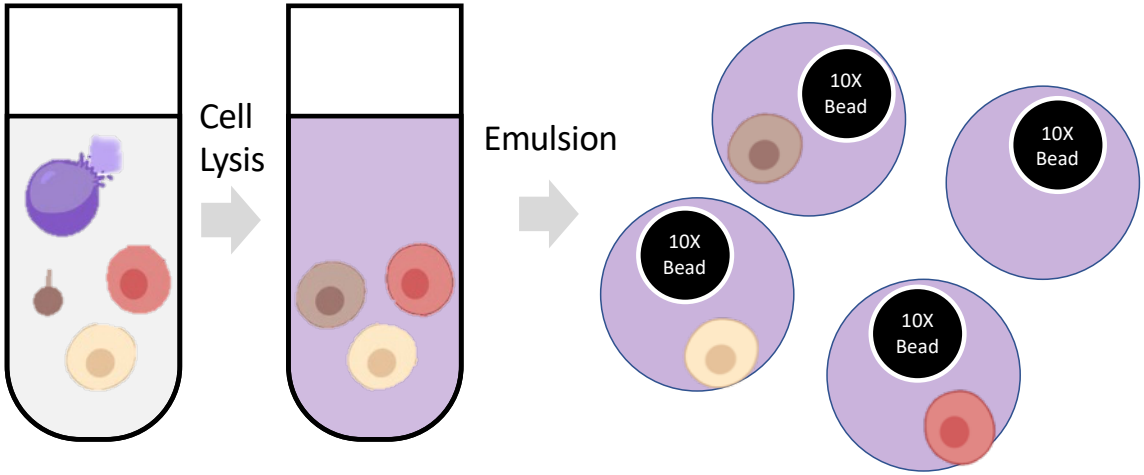
Differential expression between ”empty” droplets:

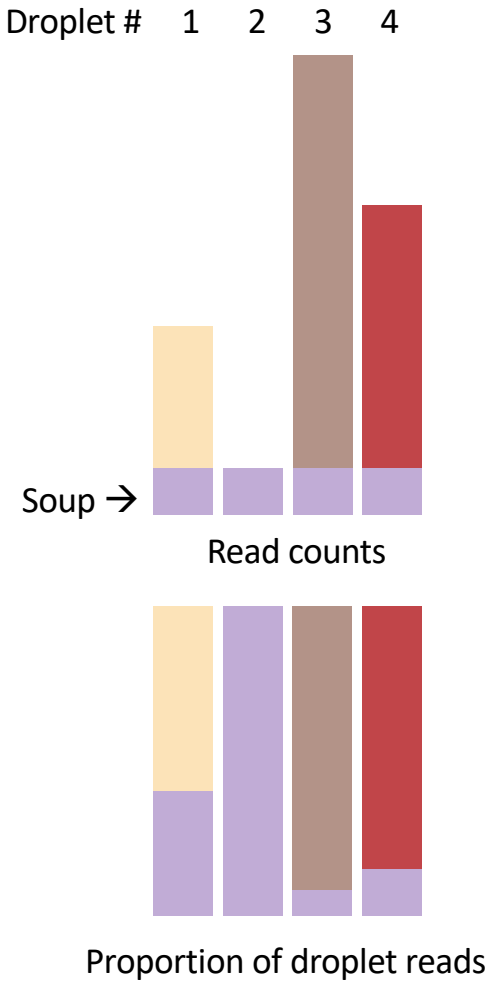


Significant sources of contaminating mRNA:

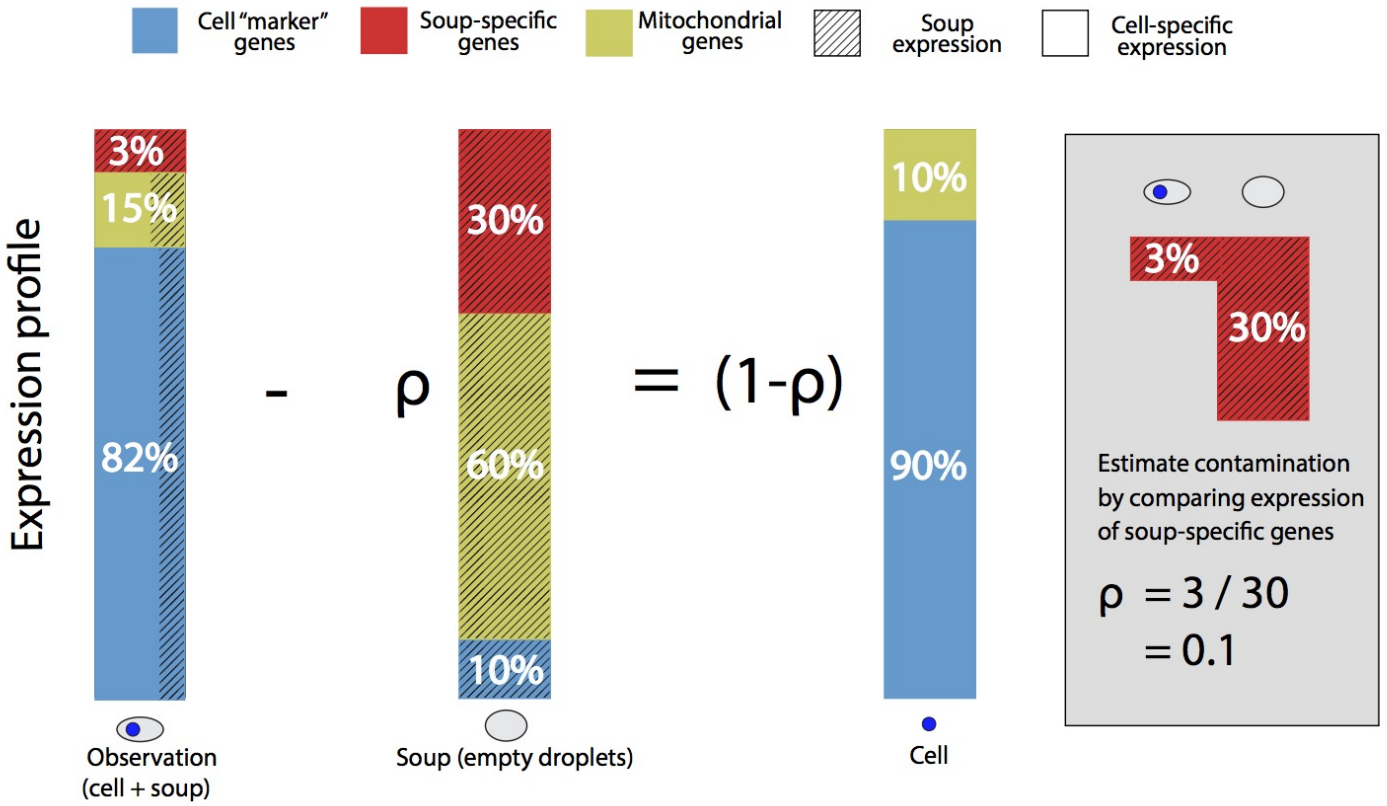
WT:	Erythrocytes
	Epithelial
	Granulocytes
KO:	Acinar cells

Ambient RNA: ”SOUP”





SoupX



Young and Behati (bioRxiv) 2018. https://www.biorxiv.org/content/10.1101/303727v1

Getting started with your own analyses

Rahul Satija - NYGC

R

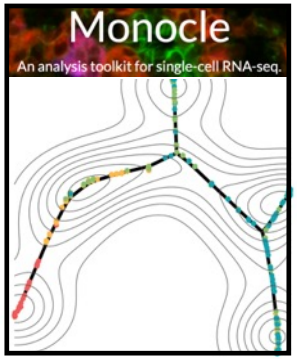
https://satijalab.org/seurat

Fabian Theis - München

https://scanpy.readthedocs.io/en/latest/

Python

R



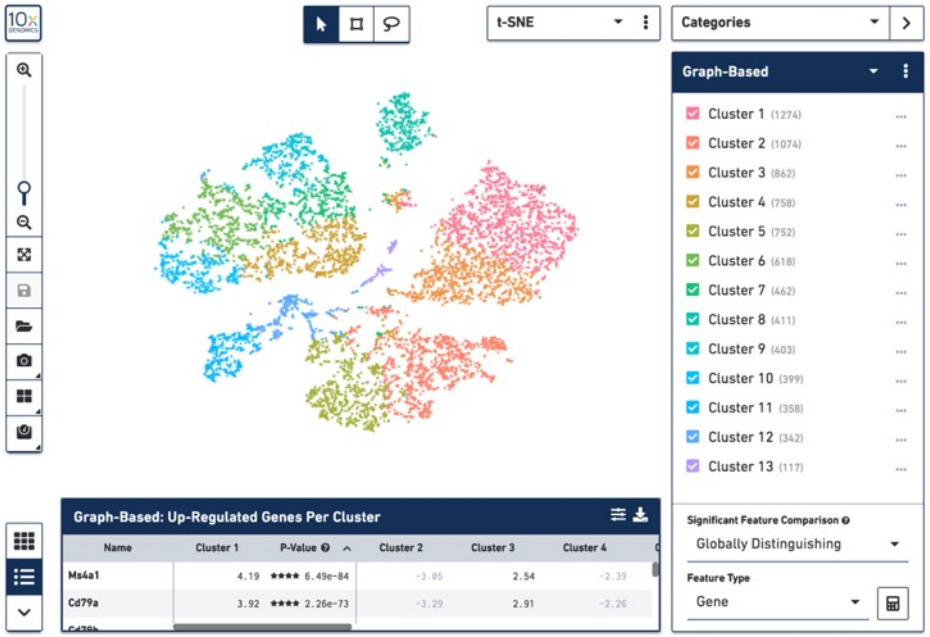
Cole Trapnell –WashU

Liger R

Macosko lab

AWESOME SINGLE CELL RESOURCE https://github.com/seandavi/awesome-single-cell

Loupe Cell Browser



https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest

Can:

- Quickly visualize genes
- Do guided clustering via marker genes / tSNE selections
- Calculate Differential Expression
- Export cells and gene sets for reanalysis on Cellranger (cluster)

Can't

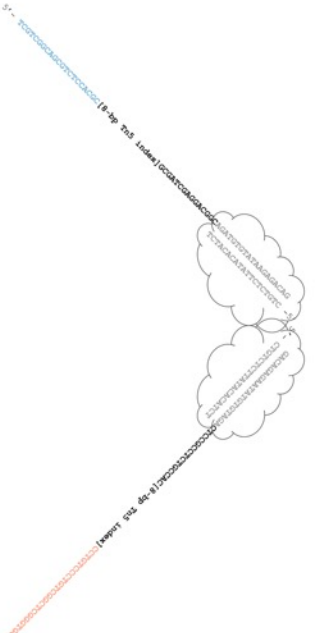
- Redo unsupervised clustering / tSNE / UMAP
- Repeat PCA / gene set selection
- Pseudotime, other fancy things

The Best Site On the Internet.

Probably.

- https://github.com/Teichlab/scg_lib_structs

Detailed visual guides to dozens of single-cell genomics methods



Adapter and primer sequences:

Barcoded Tn5 sequence s5: 5'- TCGTCGGCAGCGTCTCCACGC[8-bp Tn5 index]GCGATCGAGGACGGCAGATGTGTATAAGAGACAG -3'

Barcoded Tn5 sequence s7: 5'- GTCTCGTGGGCTCGGCTGTCCCTGTCC[8-bp Tn5 index]CACCGTCTCCGCCTCAGATGTGTATAAGAGACAG -3'

Tn5 binding site 19-bp Mosaic End (ME) bottom: 5'- /Phos/AGATGTGTATAAGAGACAG -3'

P5 index primer entry point (s5): 5'- TCGTCGGCAGCGTCTCCACGC -3'

P7 index primer entry point (s7): 5'- GTCTCGTGGGCTCGGCTGTCCCTGTCC -3'

P5 index primer: 5'- AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTCTCCACGC -3'

P7 index primer: 5'- CAAGCAGAAGACGGCATACGAGAT[i7]GTCTCGTGGGCTCGGCTGTCCCTGTCC -3'

Read 1 sequencing primer: 5'- GCGATCGAGGACGGCAGATGTGTATAAGAGACAG -3'

Index 1 sequencing primer (i7): 5'- CTGTCTCTTATACACATCTGAGGCGGAGACGGTG -3'

Read 2 seuquencing primer: 5'- CACCGTCTCCGCCTCAGATGTGTATAAGAGACAG -3'

Product 1 (s5 at both ends, not amplifiable due to semi-suppressive PCR:

5'- TCGTCGGCAGCGTCTCCACGC[8-bp Tn5 index]GCGATCGAGGACGGCAGATGTGTATAAGAGACAGXXXXXXXXXXXX...XXX CTGTCTCTTATACACATCT
TCTACACATATTCTCTGTC XXX...XXXXXXXXXXXXGACAGAGAATATGTGTAGACGGCAGGAGCTAGCG[8-bp Tn5 index]CGCACCTCTGCGACGGCTGCT -5'

Product 2 (s7 at both ends, not amplifiable due to semi-suppressiev PCR):

5'- GTCTCGTGGGCTCGGCTGTCCCTGTCC[8-bp Tn5 index]CACCGTCTCCGCCTCAGATGTGTATAAGAGACAGXXXXXXXXXXXX...XXX CTGTCTCTTATACACATCT
TCTACACATATTCTCTGTC XXX...XXXXXXXXXXXXGACAGAGAATATGTGTAGACTCCGCCTCTGCCAC[8-bp Tn5 index]CCTGTCCCTGTCGGCTCGGGTGCTCTG -5'

"What I cannot create, I do not understand." --Feynman

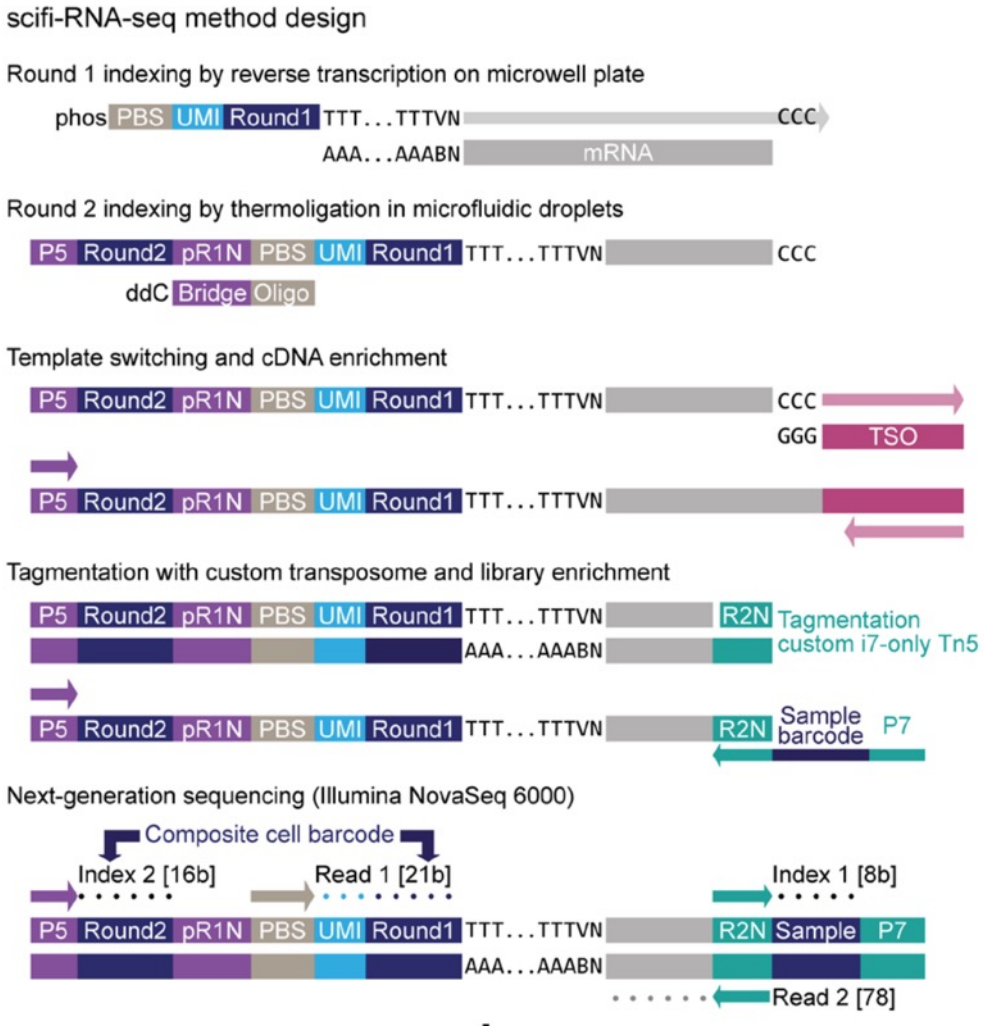
Hacking Droplets

scifi-RNA-seq

Combinatorial fluidic indexing

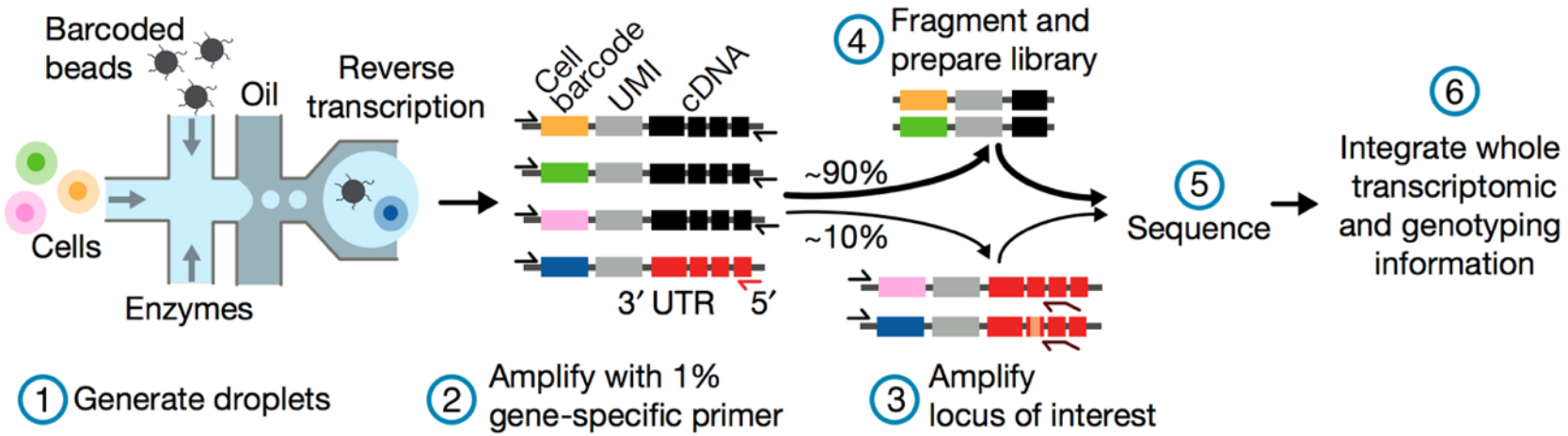
- Massive improvement in # cells
- Up-front barcoding in plates via RT
- Swaps chemistry of 10X Genomics:
  - Uses 10X Gel beads
  - Ligation instead of RT

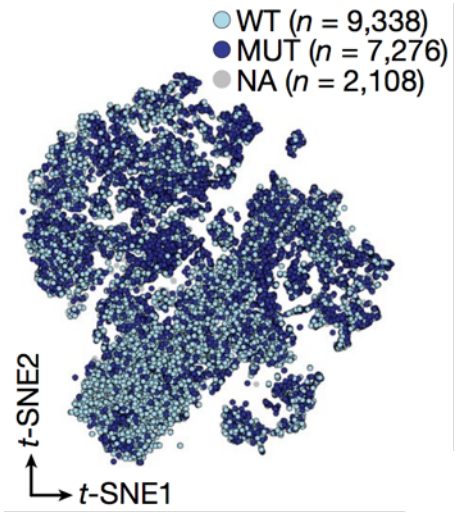
- Up to 150,000 cells per channel
  - (15X increase)

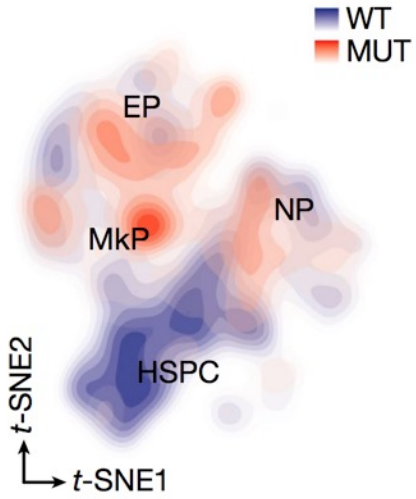


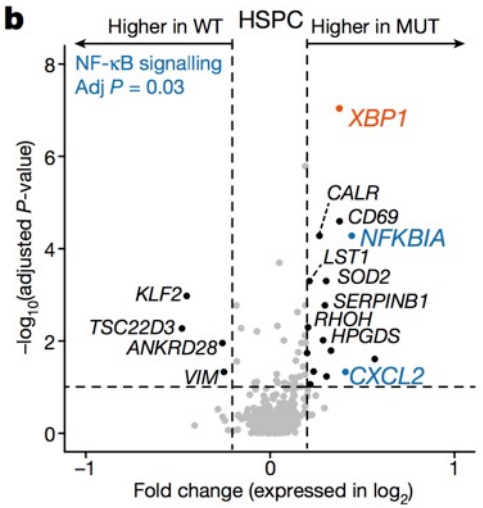
Paul Datlinger, André F Rendeiro, Thorina Boenke, Thomas Krausgruber, Daniele Barreca, Christoph Bock
doi: https://doi.org/10.1101/2019.12.17.879304

Genotyping of Transcriptomes



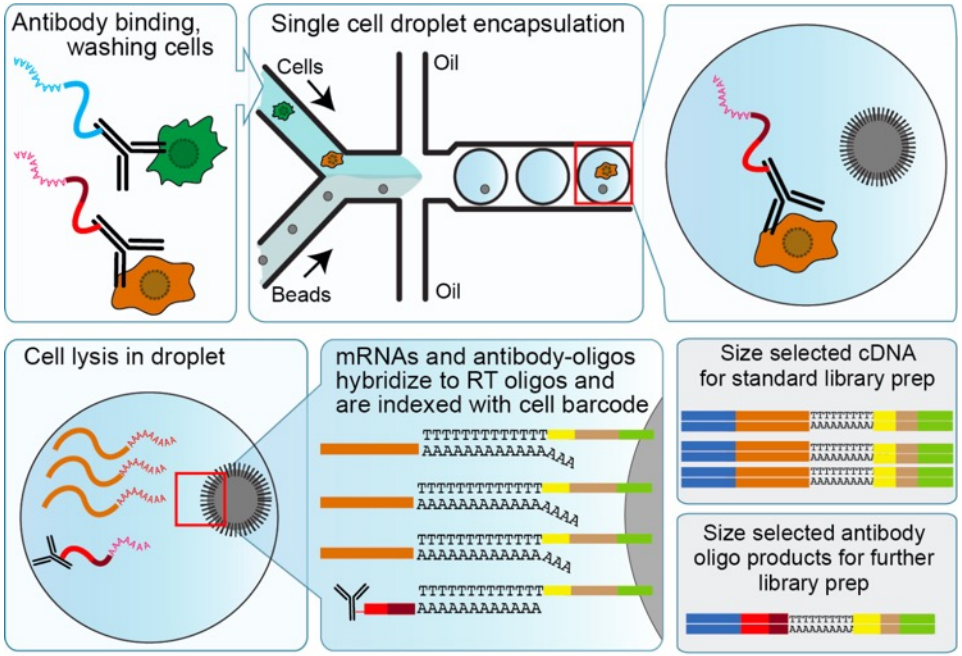






Nam et al. Nature **571**, 355–360(2019)

CITE-Seq / REAP-Seq



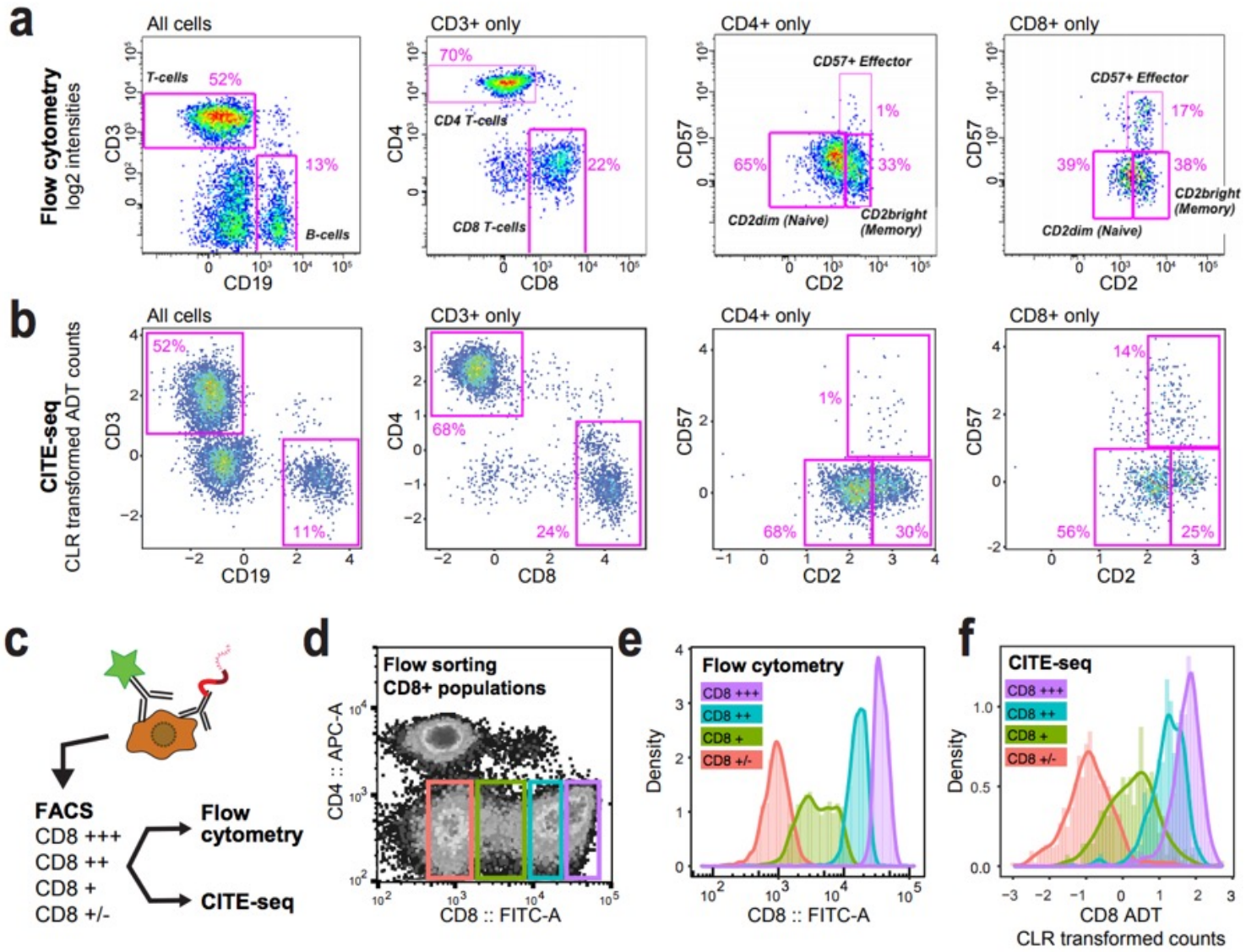
Antibody **D**erived **T**ag (**ADT**)
sequenced as part of normal 10X run

Enables:

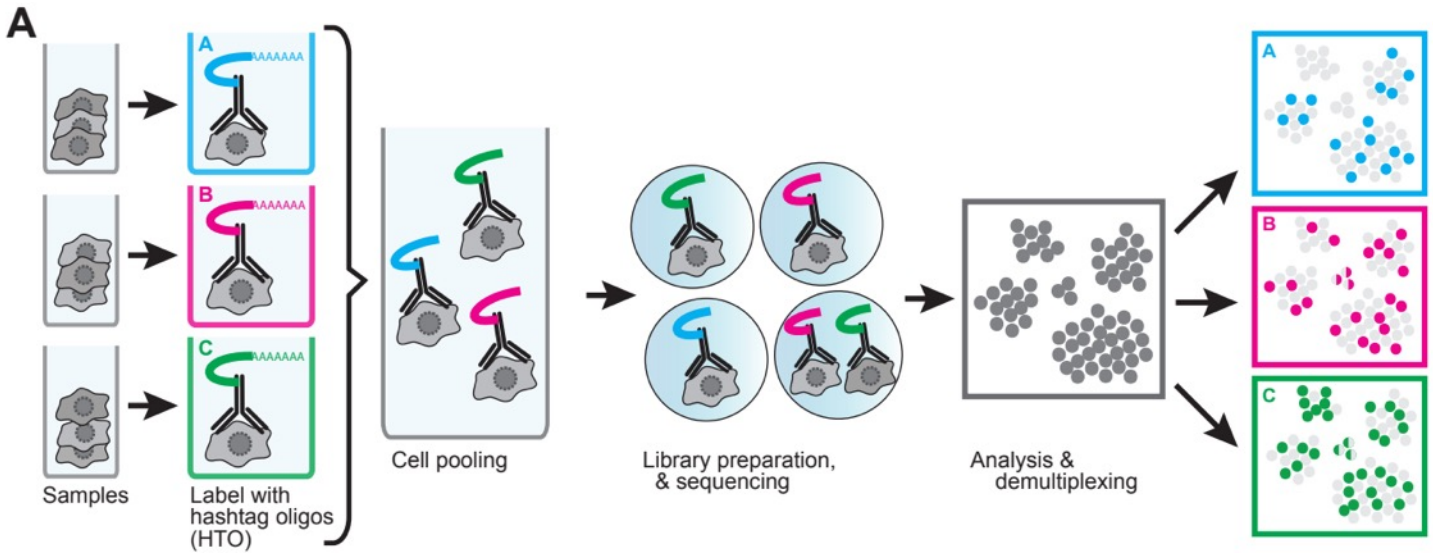
- Simultaneous mRNA + Protein Abundance
- Increased sensitivity to individual targets
- 'Superloading'

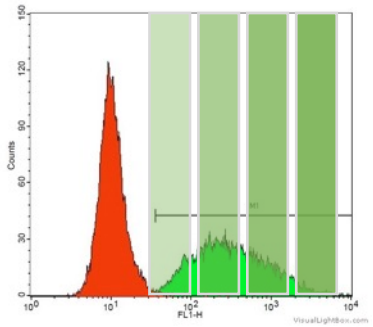
Stoeckius et al. *Nature Methods* **14**, 865–868 (2017)

CITE-Seq / REAP-Seq



Multiplexing with ADTs: “Cell Hashing”

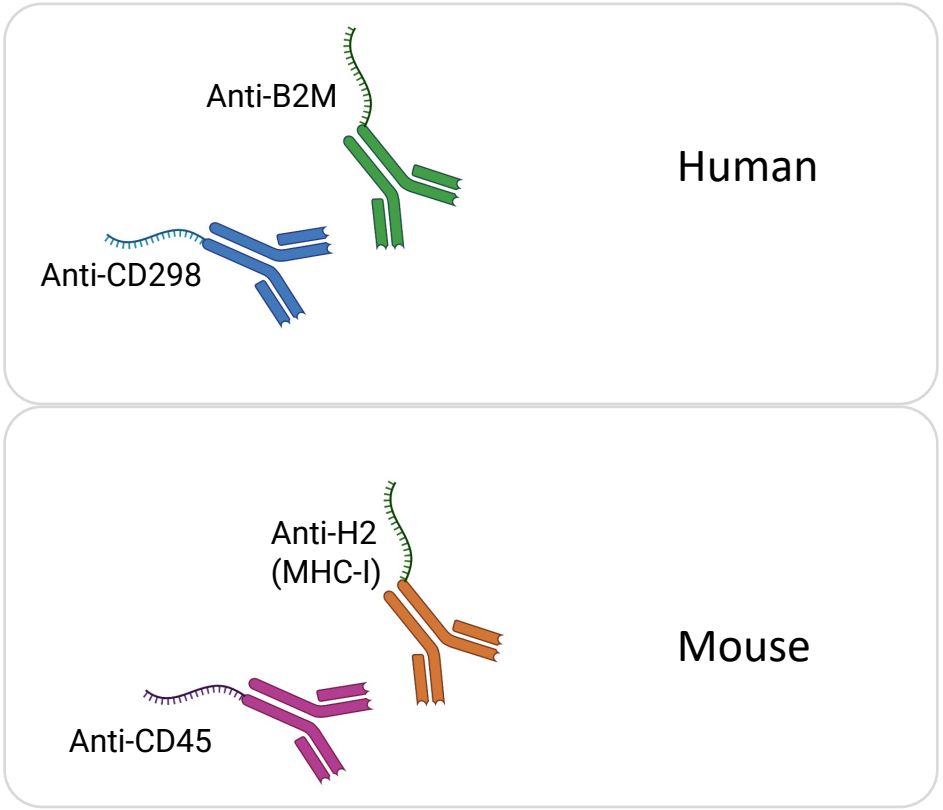




Sort multiple bins → HTO Label → Repool & Capture

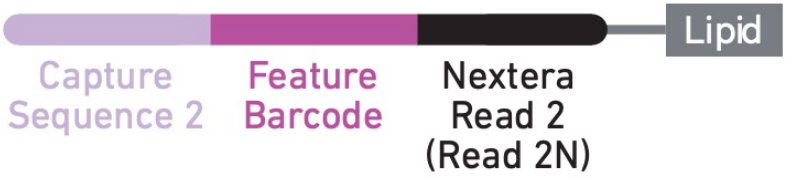
Stoeckius, NYGC

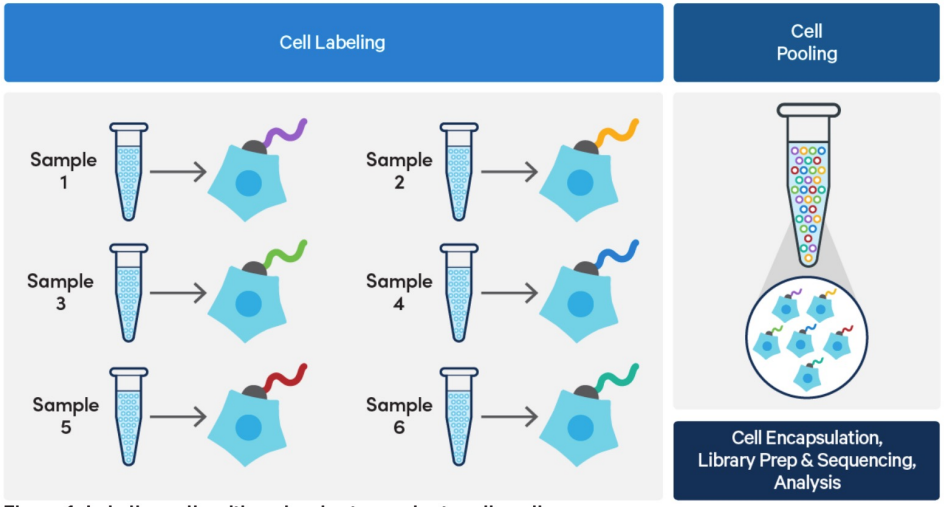
Total-Seq (Biolegend)

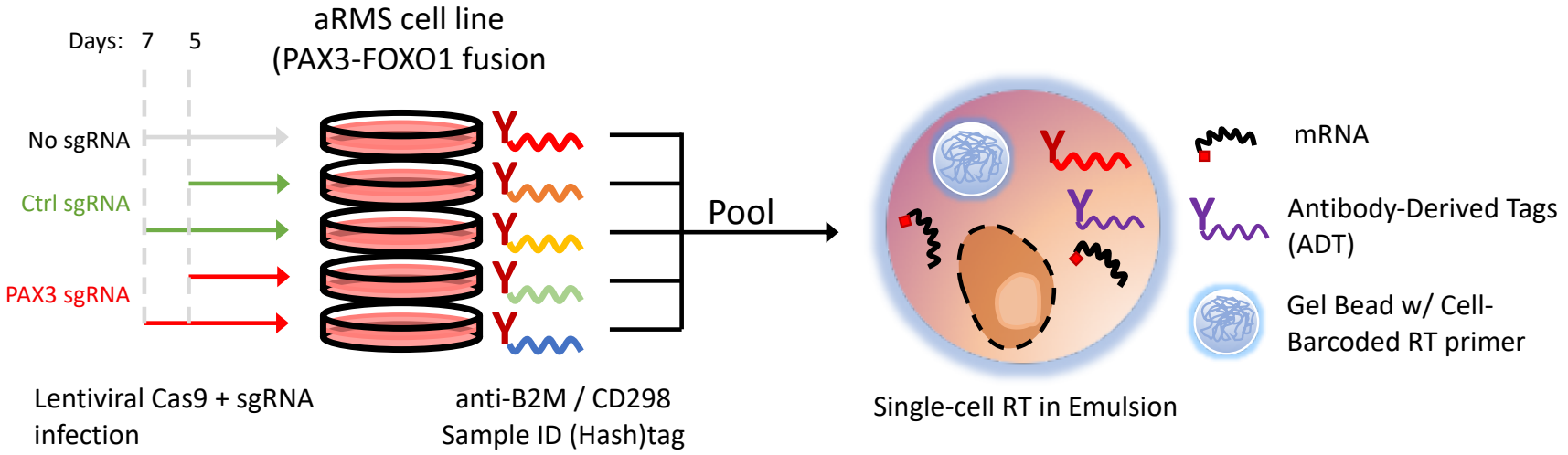


CellPlex (10X Genomics)

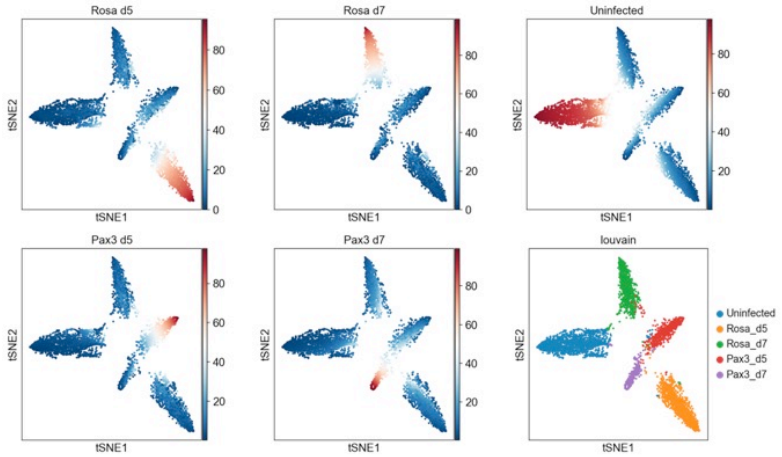
Cholesterol / Lipid anchor



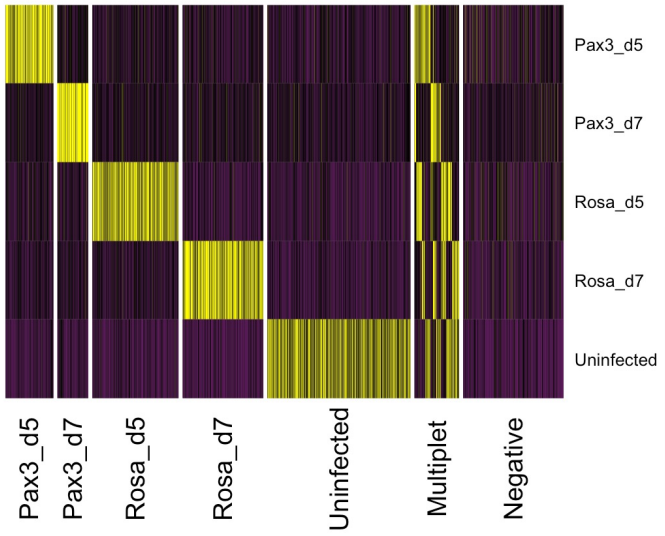


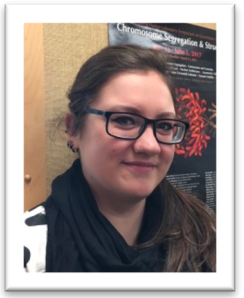


tSNE by Hash Tag

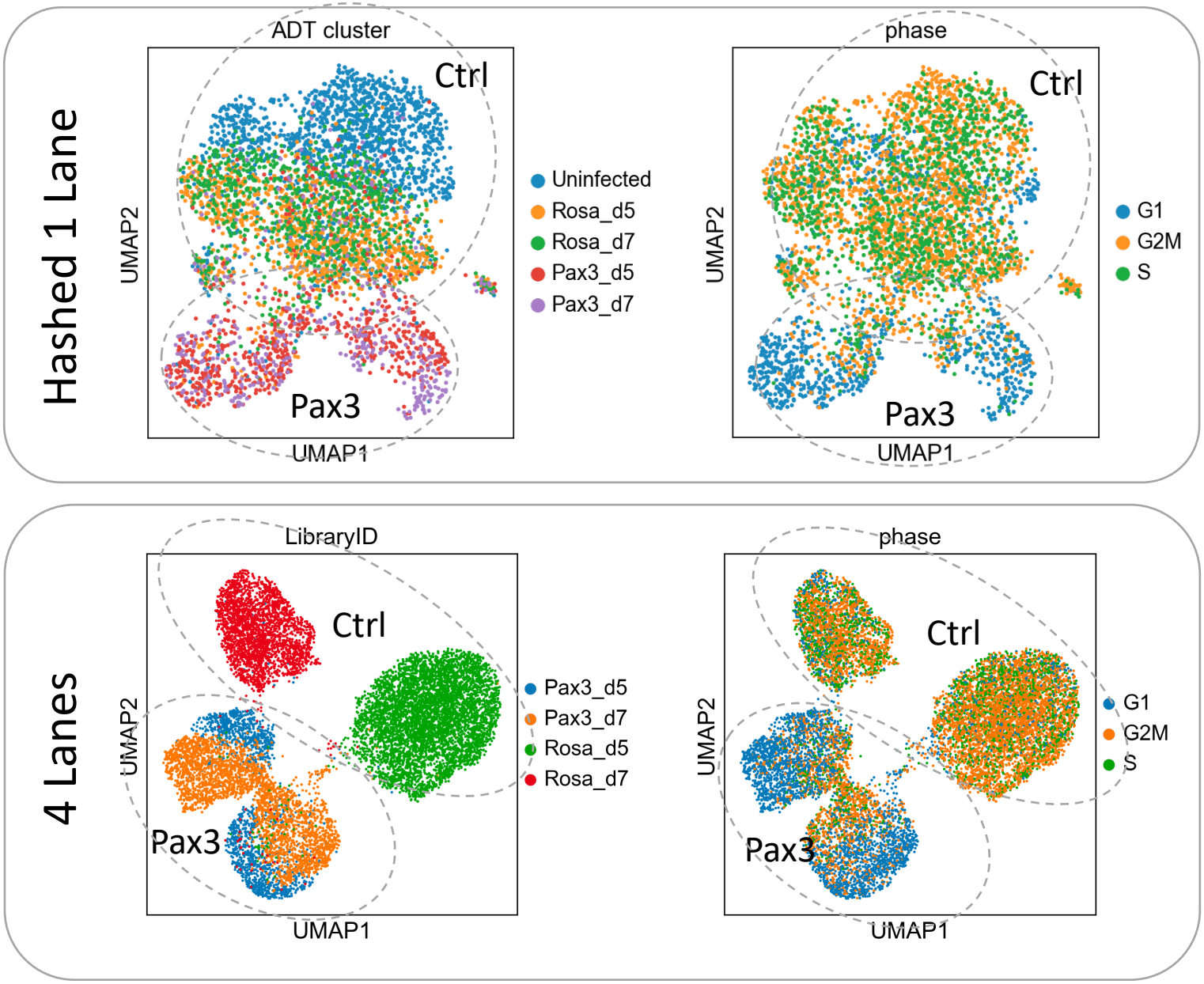


Hash tag Clustering

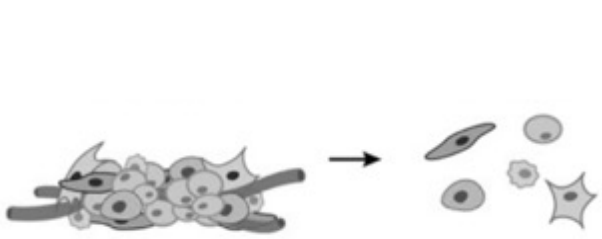




Martyna Sroka, Vakoc Lab

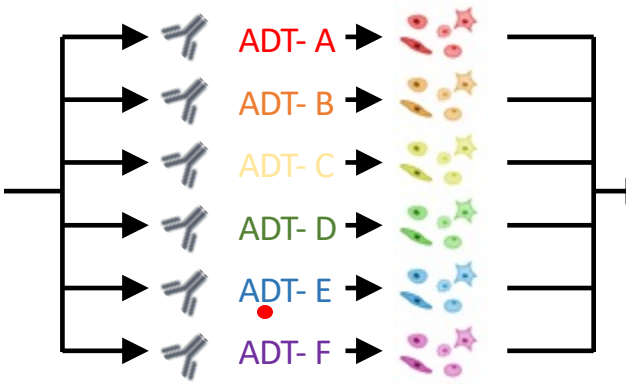

Hashed 1 Lane
ADT cluster
Ctrl
Pax3
UMAP2
UMAP1
Uninfected
Rosa_d5
Rosa_d7
Pax3_d5
Pax3_d7
phase
Ctrl
Pax3
UMAP2
UMAP1
G1
G2M
S
4 Lanes
LibraryID
Ctrl
Pax3
UMAP2
UMAP1
Pax3_d5
Pax3_d7
Rosa_d5
Rosa_d7
phase
Ctrl
Pax3
UMAP2
UMAP1
G1
G2M
S

Superloading with ADTs: “Cell Hashing”

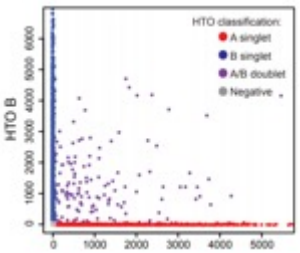


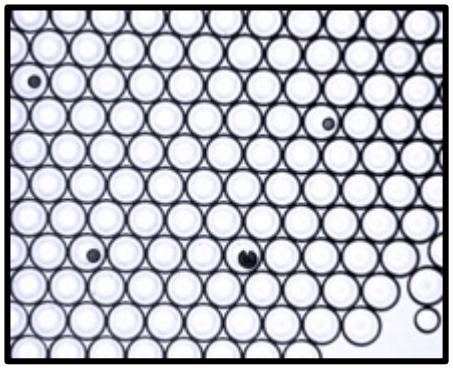
Universal surface marker:

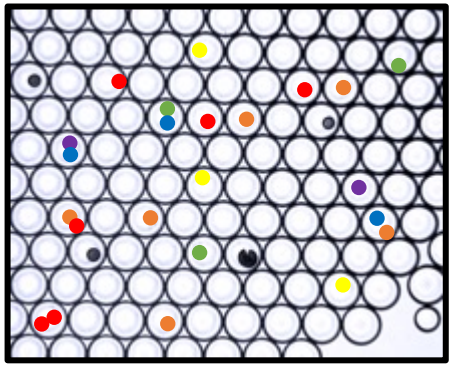
CD98 (neutral amino acid transporter)
B2M

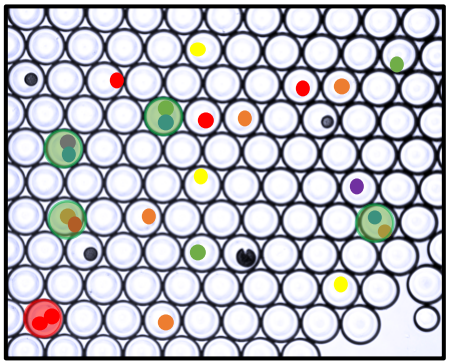


Identify single cells by hash







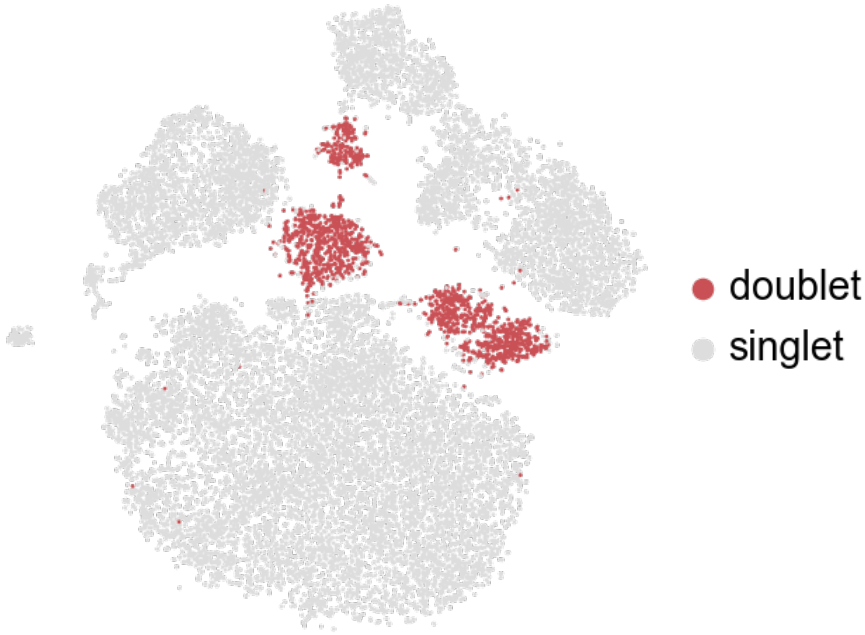


Filter Multiplets
in silico

Visible Doublet
(2 barcodes)

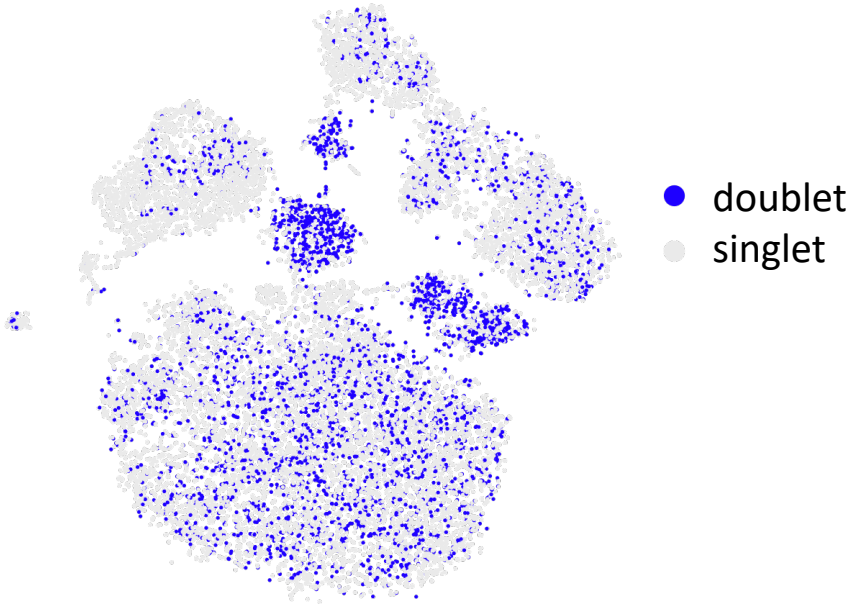
Invisible Doublet
(same barcode)

Doublet Detection by Cell Cluster



singlet	17,110
doublet	1,668

Doublet Detection by Hash Tag

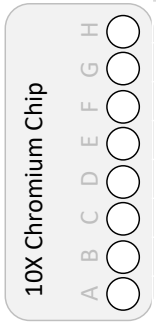


singlet	15,148
doublet	3,630

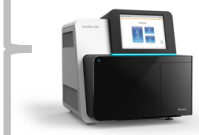
Superloading with Cell Hashing Benefits

Setup Fee: $500

$/Channel	Cells
$1400	6,000
$1400	6,000
$1400	6,000
$1400	6,000
$1400	6,000
$1400	6,000
$1400	6,000
$1400	6,000



NextSeq HO x4



Cost Per Cell:

Capture: $0.24
Sequencing: $0.15

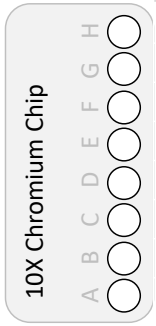
Total Experiment Cost:

Total Capture: $11,700
Total Sequencing: $7,200

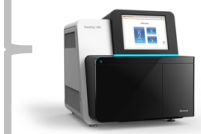
Total: $18,900

Setup Fee: $500

$/Channel	Cells
$1400	24,000
$1400	24,000



NextSeq HO x4



Cost Per Cell:

Capture: $0.07
Sequencing: $0.15

Total Experiment Cost:

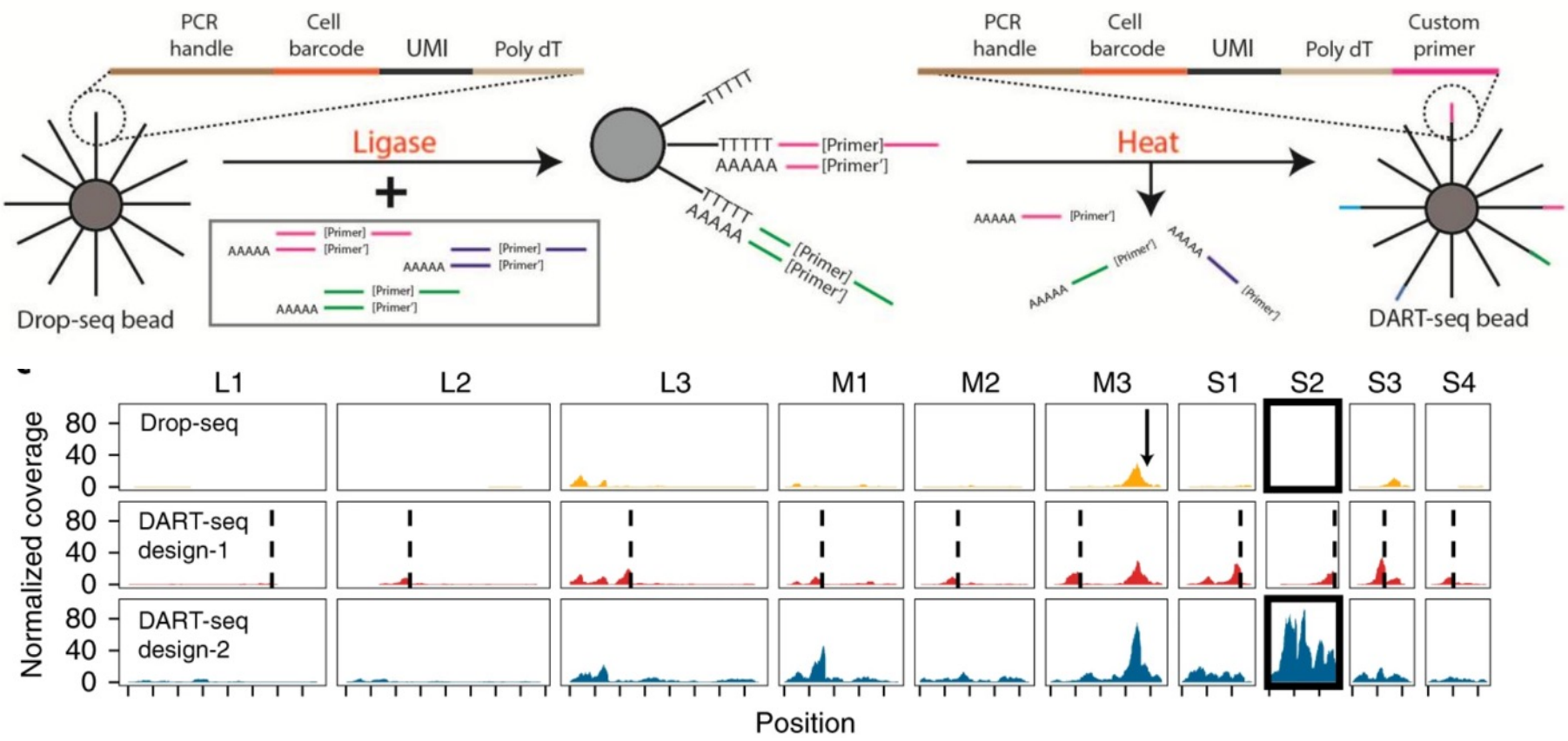
Total Capture: $3,300
Total Sequencing: $7,200

Total: $10,500

DART-seq

Droplet-Assisted RNA Targeting by single-cell sequencing

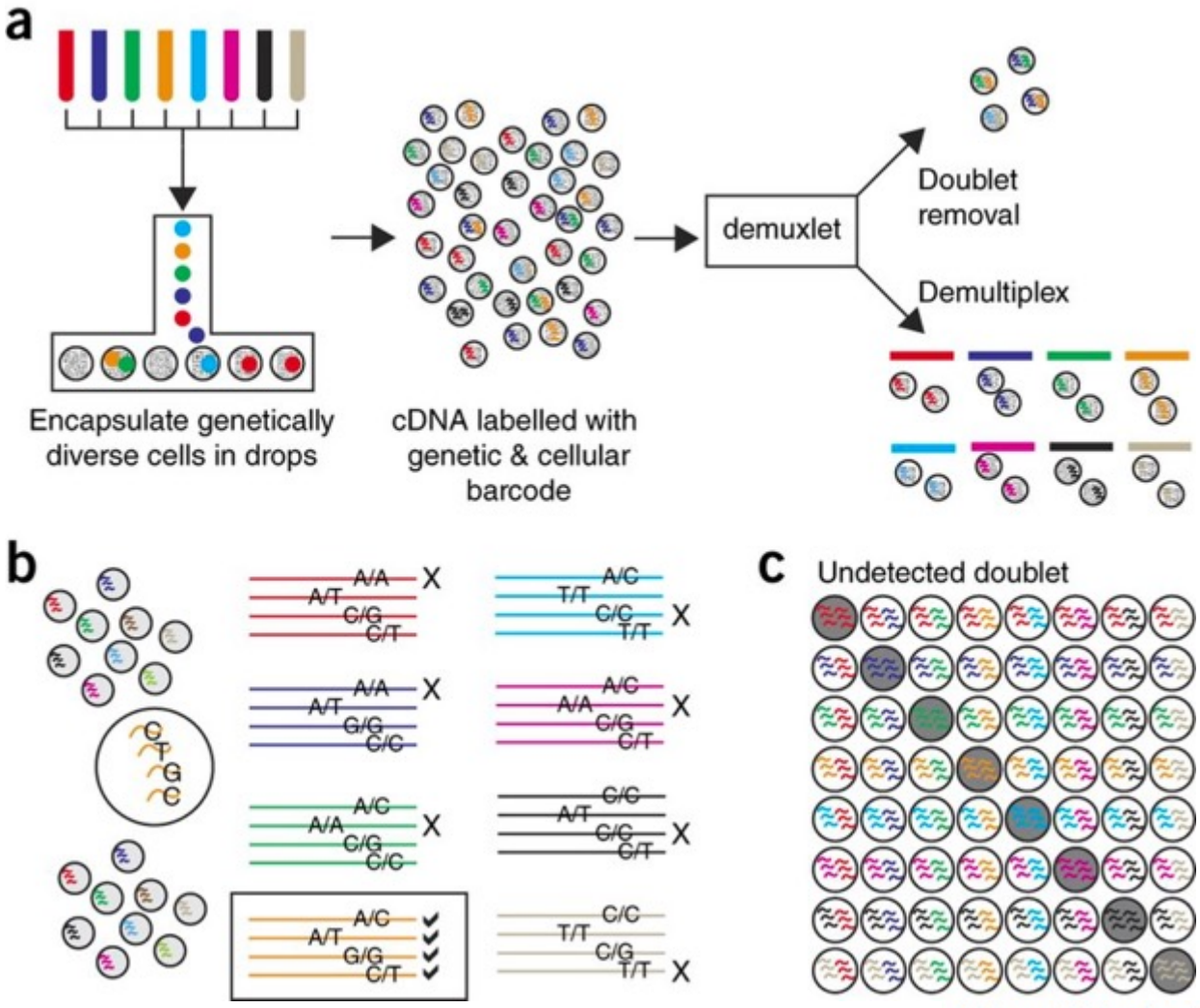
- Modification of barcoded bead to prime non-poly(A) transcripts
- Ligate gene-specific primers to subset of oligo-dT sites via bridge oligo
  - Careful titration of primers necessary



Saikia, M., Burnham, P., Keshavjee, S.H. *et al.* Simultaneous multiplexed amplicon sequencing and transcriptome profiling in single cells. *Nat Methods* **16,** 59–62 (2019). https://doi.org/10.1038/s41592-018-0259-9

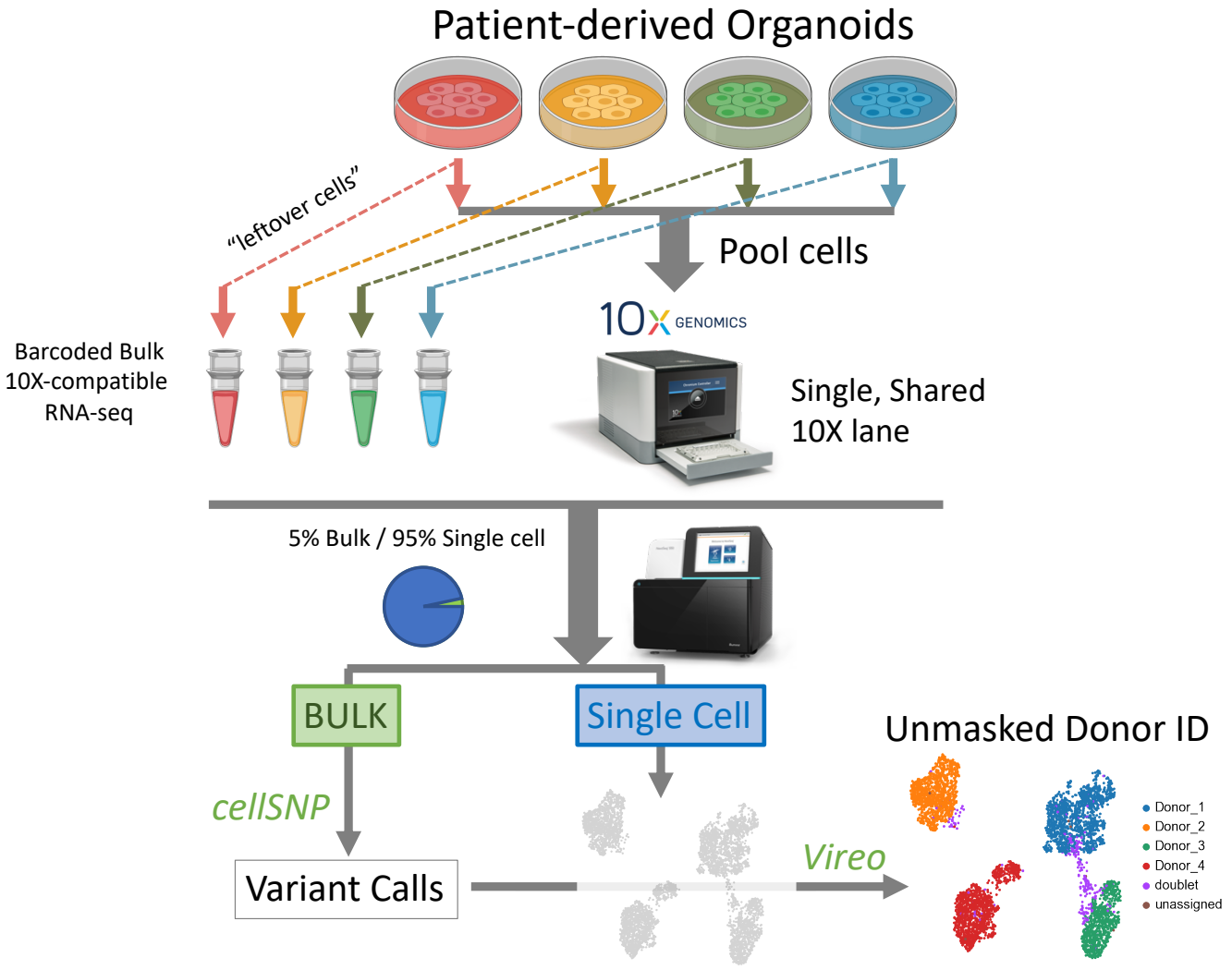
Multiplexing Using Natural Genetic Variation

Demuxlet

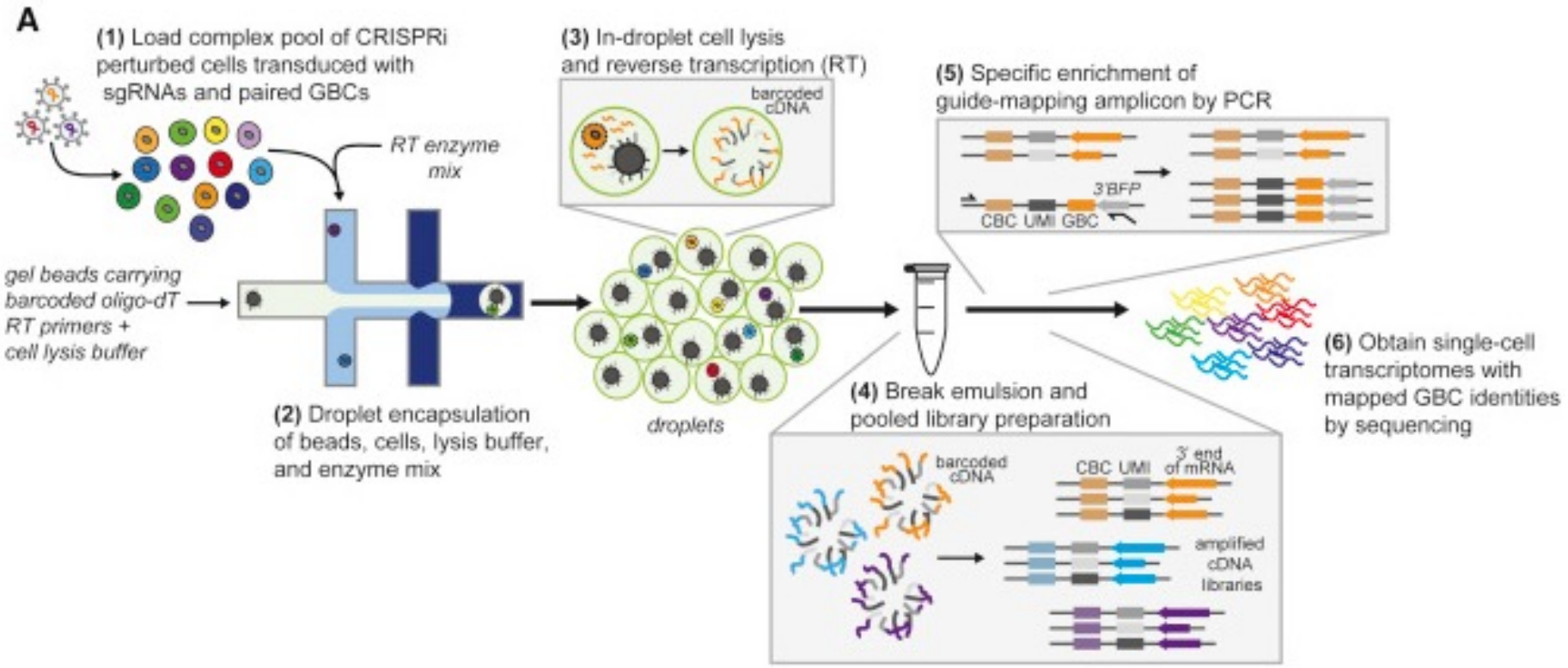


Jimmie Ye lab

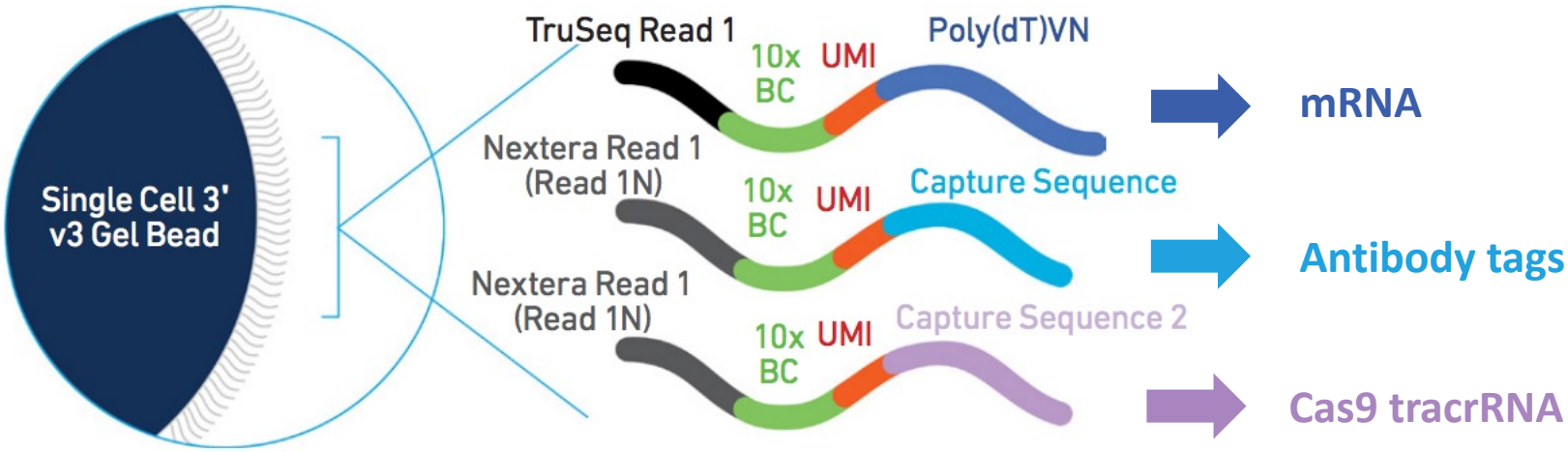
Nature Biotechnology **36**, 89–94 (2018)


Patient-derived Organoids
"leftover cells"
Pool cells
Barcoded Bulk
10X-compatible
RNA-seq
10X GENOMICS
Single, Shared
10X lane
5% Bulk / 95% Single cell
BULK
Single Cell
cellSNP
Variant Calls
Vireo
Unmasked Donor ID
Donor_1
Donor_2
Donor_3
Donor_4
doublet
unassigned

Perturb-Seq



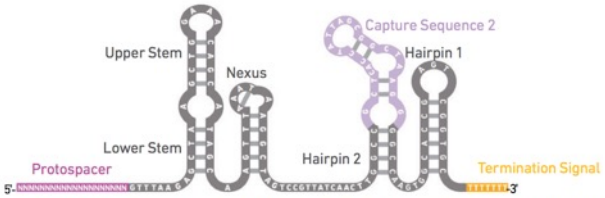
Adamson et al. (2016). Cell 16(7), 1867–1882.e21



- Multiple RT primer sequences per bead
- High efficiency capture of antibody tags, CRISPR guides

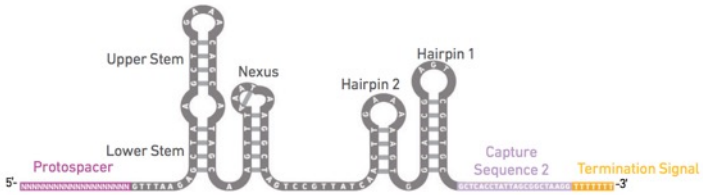
Capture Sequence 2 on Gel Bead: 5'-CCTTAGCCGCTAATAGGTGAGC-3'

Capture Sequence 2 integrated in sgRNA hairpin



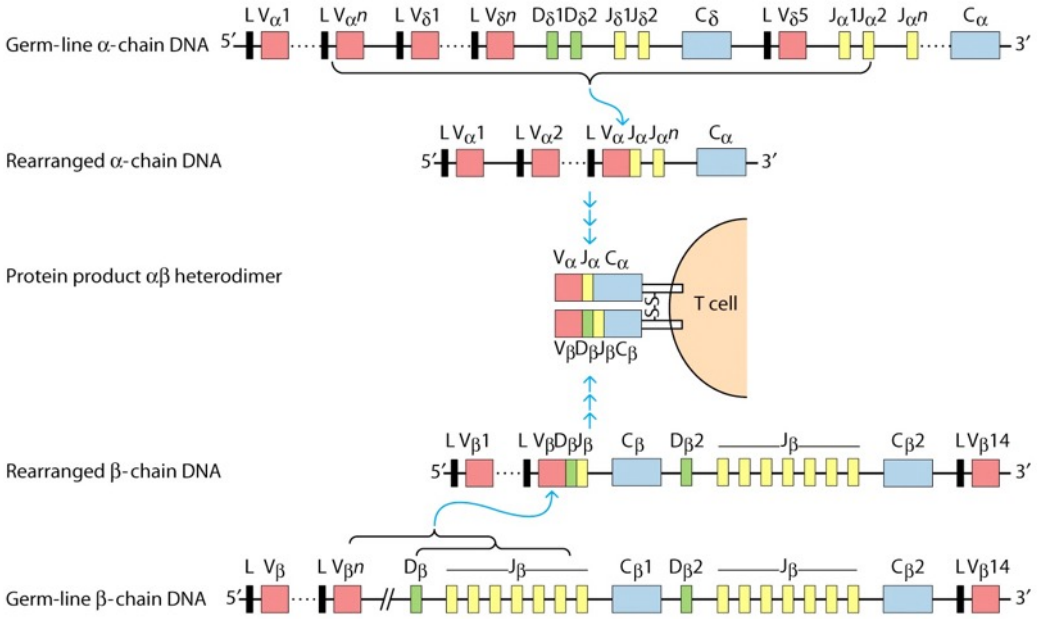
5'-NNNNNNNNNNNNNNNNNNNGTTTAAGAGCTAAGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTggccGCTCACCTATTAGCGGCTAAGGggccAAGTGGCACCGAGTCGGTGCTTTTTTT-3'

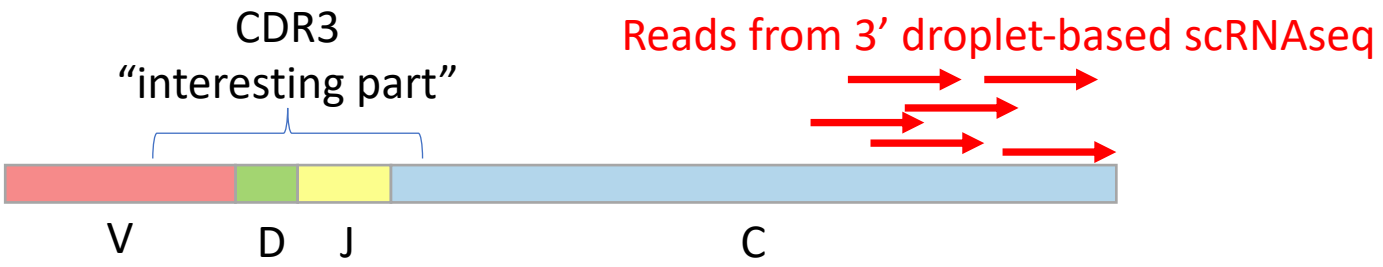
Capture Sequence 2 integrated in sgRNA 3'-end



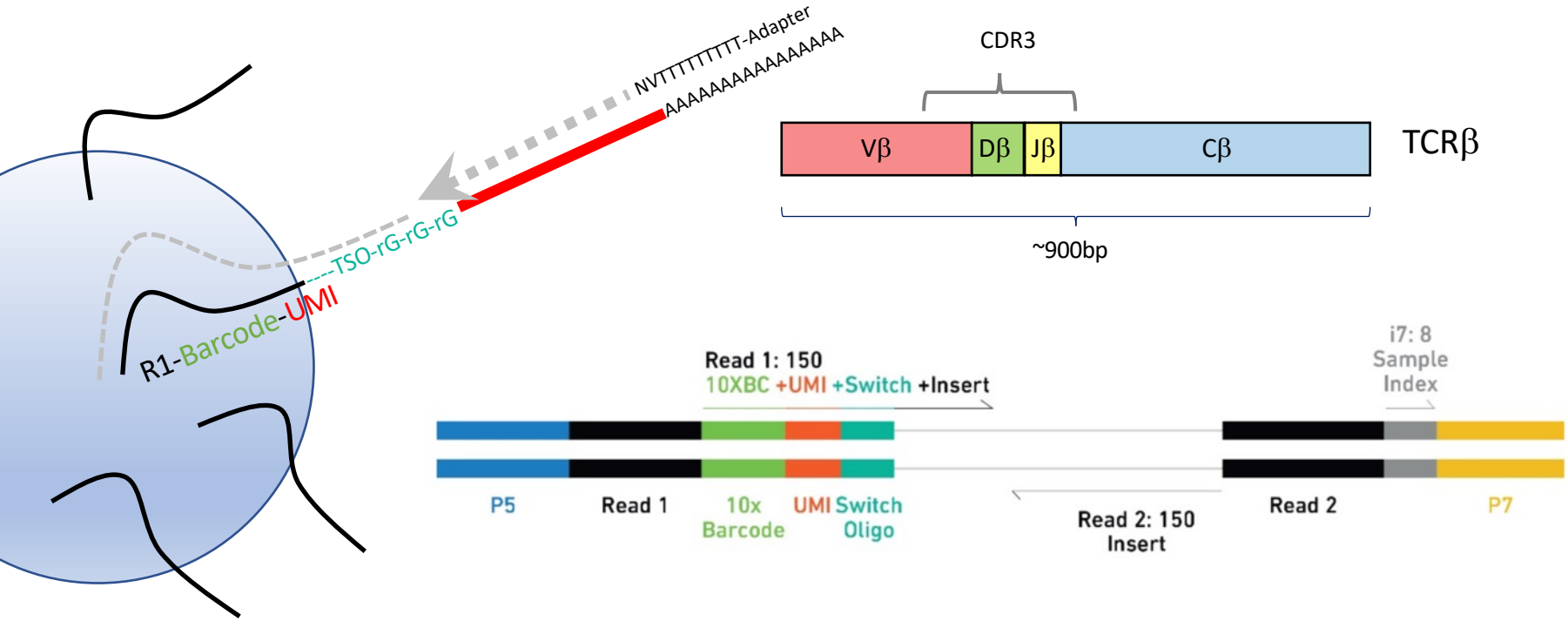
5'-NNNNNNNNNNNNNNNNNNNGTTTAAGAGCTAAGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTgaaaAAGTGGCACCGAGTCGGTGCGCTCACCTATTAGCGGCTAAGGTTTTTTT-3'

TCR/BCR Profiling





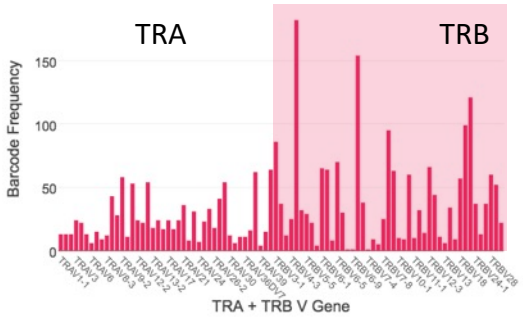
5’-Barcoded Libraries

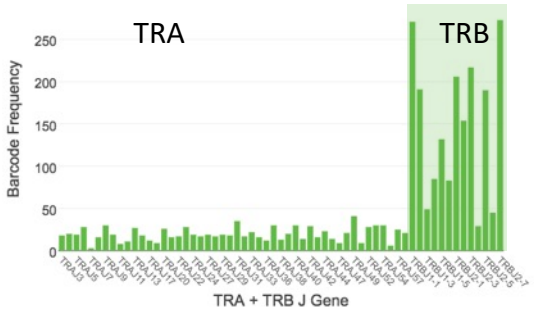


- Problem: standard transcriptome libraries have strong 3’-bias
- CDR3 mapping requires 5’-Barcoded library
- Random fragmentation to sample different 3’-ends of reads
- Require much longer reads (300bp) at a depth of 5,000X / cell

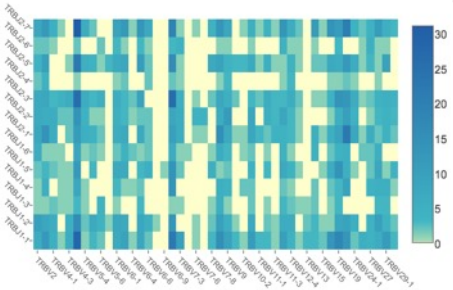
10X VDJ output example

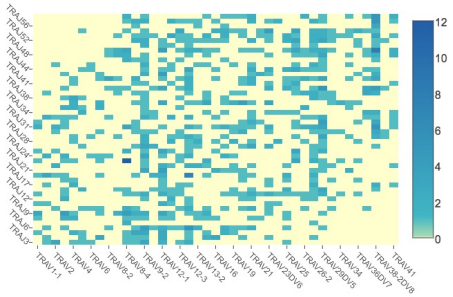
Individual V / J usage

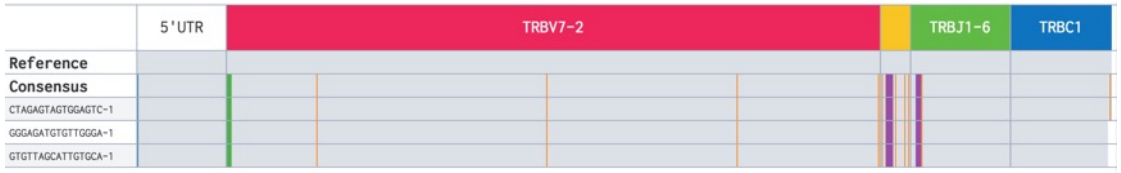




Pairwise V-J usage







Independent T-cells

Example Rearranged TCR beta chain

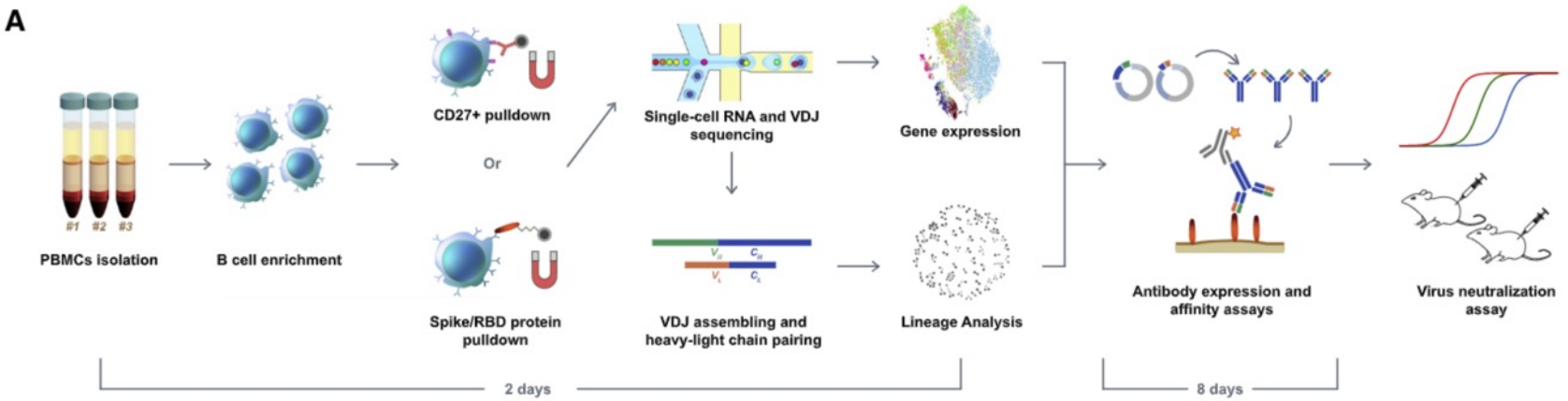
CDR3

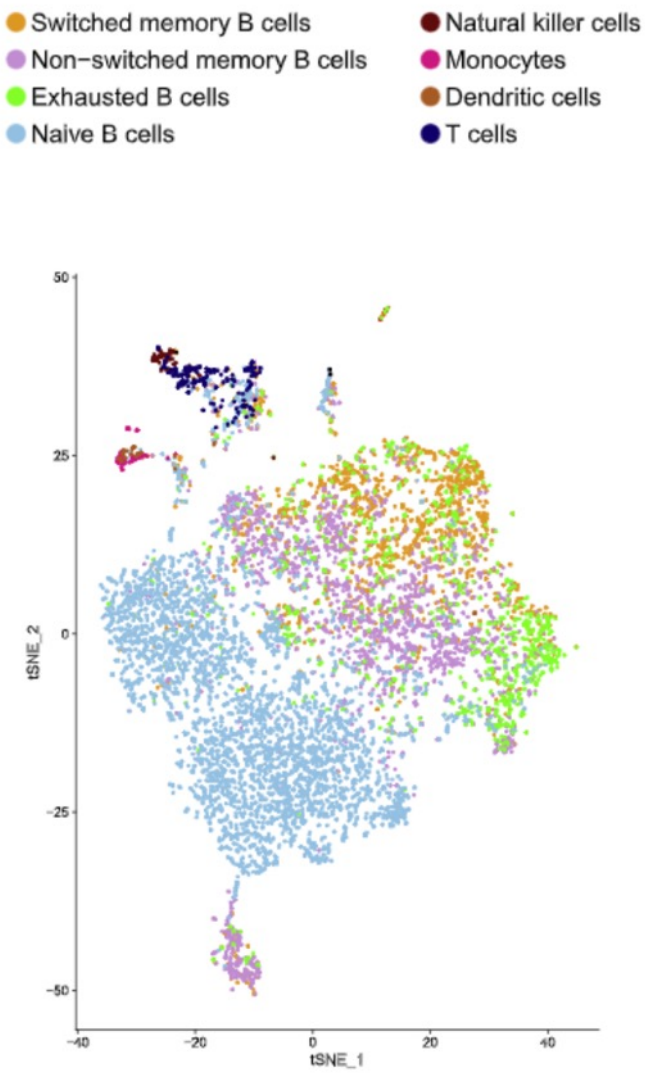
AA: CASRRGGGKTYEQYF

NT: TGTGCCAGCCGCCGGGGCGGGGGGAAAACCTACGAGCAGTACTTC

Cell

Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells





Rapid Filter for Neutralizing antibody candidates:

VDJ sequencing:

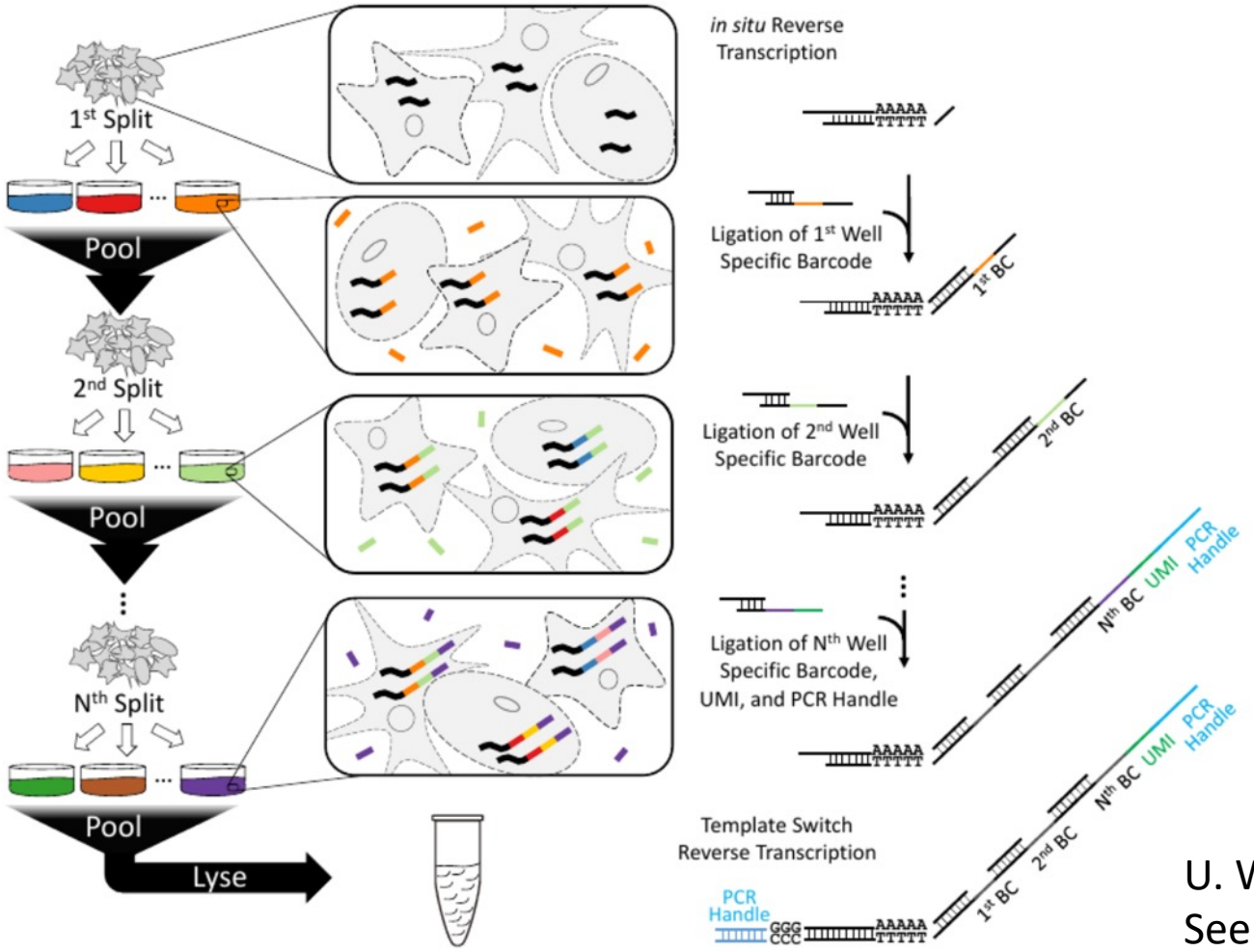
1. Select only IgG1 isotypes
2. Clones with multiple observed cells
3. Clones with somatic hypermutation

Gene expression analysis

1. Exclude exhausted and naïve phenotypes
2. Favor memory and plasma phenotypes

Other high-throughput platforms

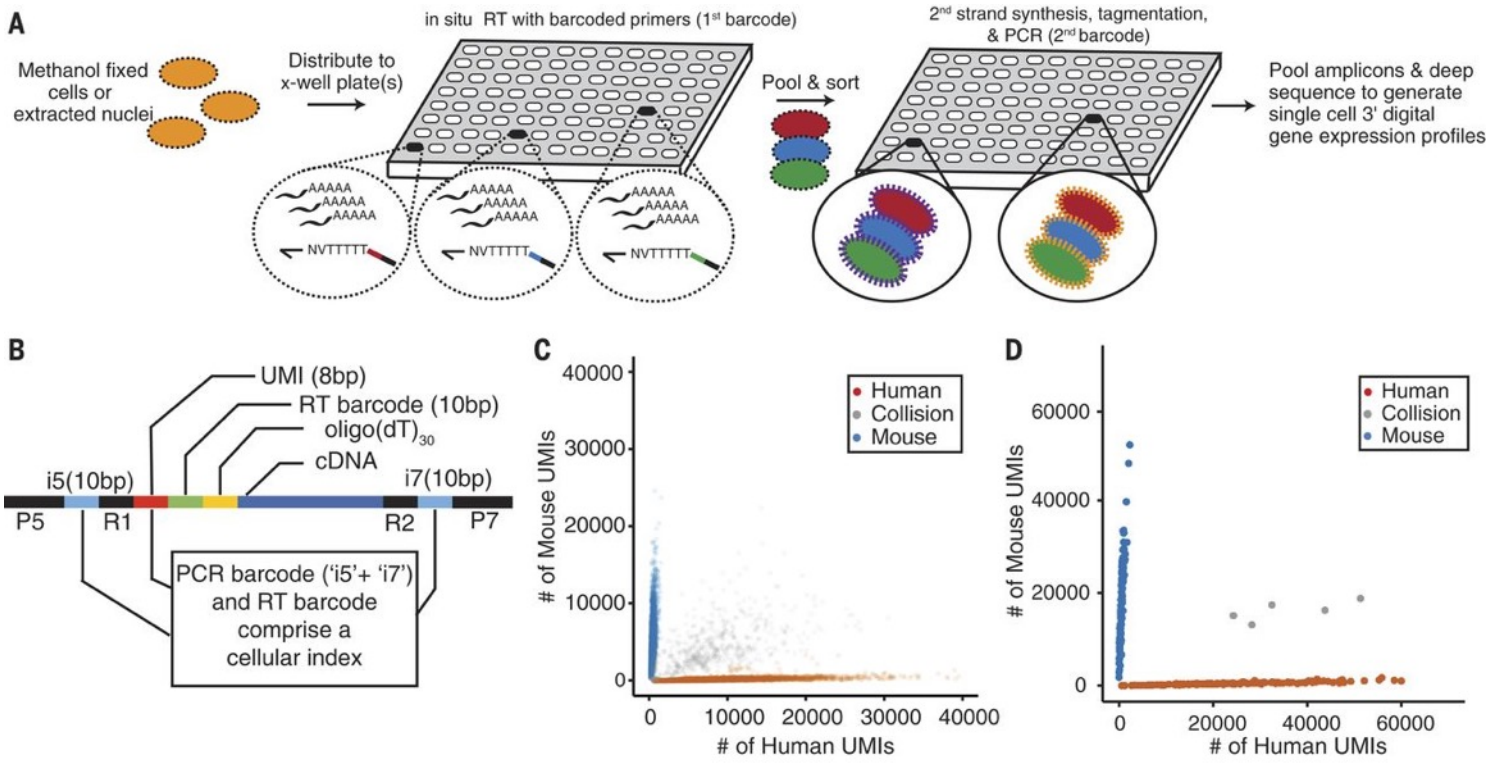
Combinatorial Indexing
Split-Seq



U. Washington
Seelig Lab

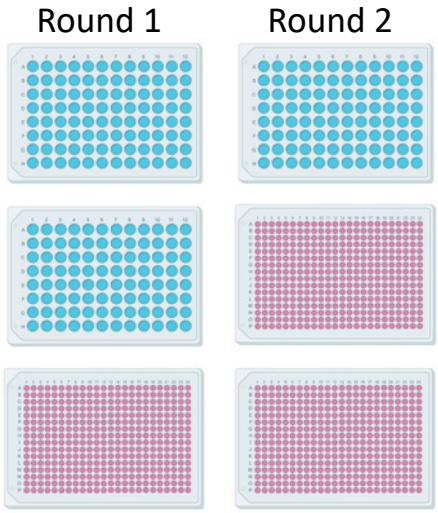
Combinatorial Indexing

sci-Seq



U. Washington
Shendure Lab

Combinatorial Scaling

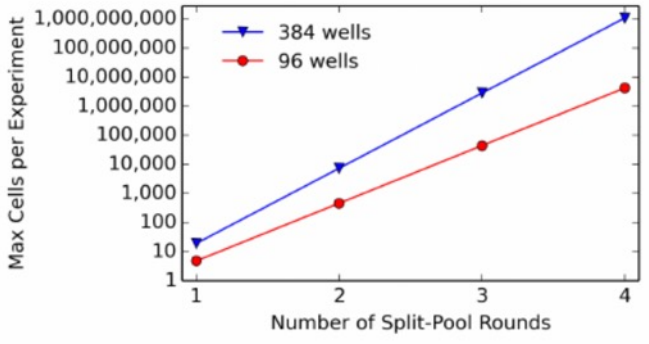


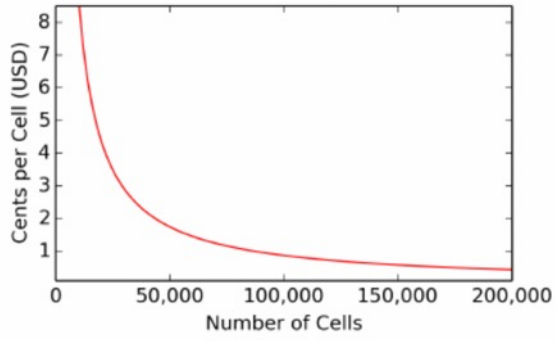
96 x 96 = 9,216 ÷ 10 = ~921 cells

96 x 384 = 36,864 ÷ 10 = ~3,684 cells

384 x 384 = 147,456 ÷ 10 = 14,746 cells

- To avoid random sampling of same barcode combinations, use ~10% of total theoretical combinations as input





Enormously scalable
Can achieve <$0.01 per cell

Labor intensive
Significant 'boot-up' cost
Significant validation cost
Who can afford that much sequencing, anyway?

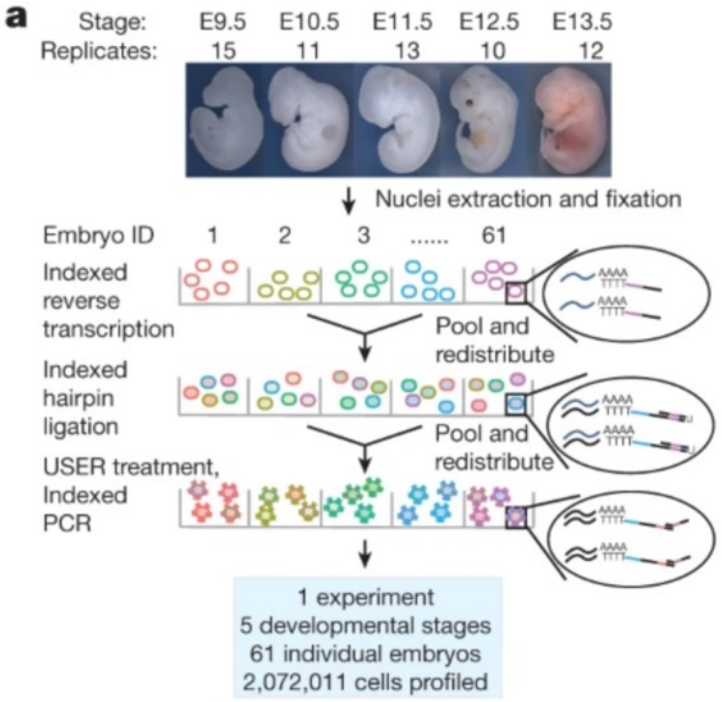
Article | Published: 20 February 2019

The single-cell transcriptional landscape of mammalian organogenesis

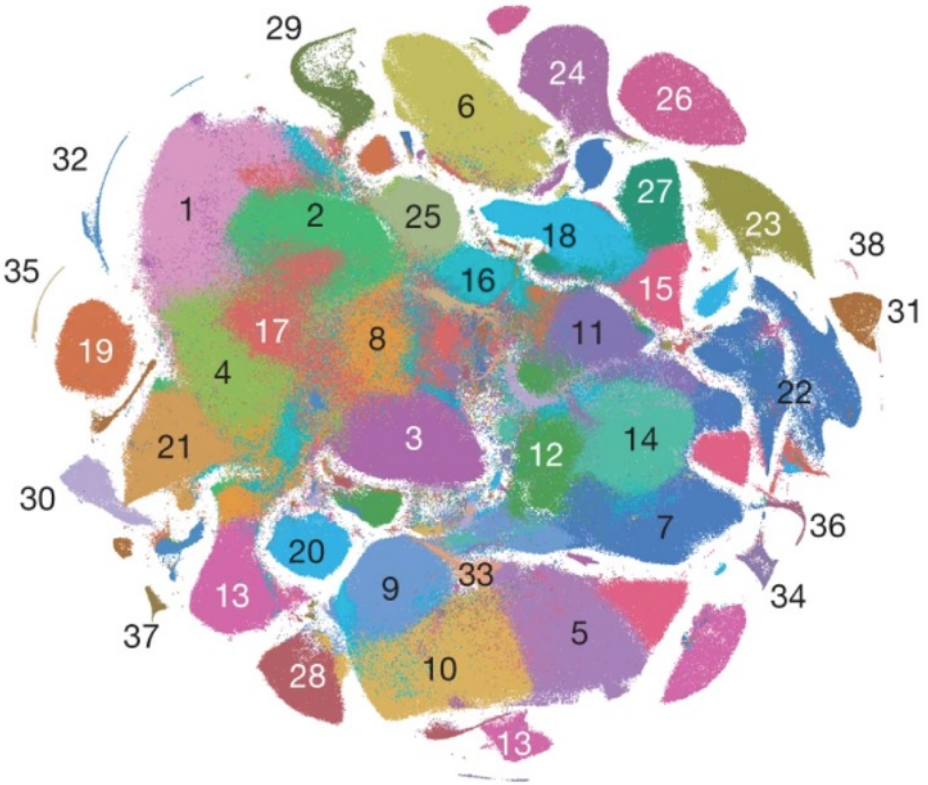
Junyue Cao, Malte Spielmann, Xiaojie Qiu, Xingfan Huang, Daniel M. Ibrahim, Andrew J. Hill, Fan Zhang, Stefan Mundlos, Lena Christiansen, Frank J. Steemers, Cole Trapnell ✉ & Jay Shendure ✉

Nature **566**, 496–502(2019) | Cite this article

sci-RNA-seq3



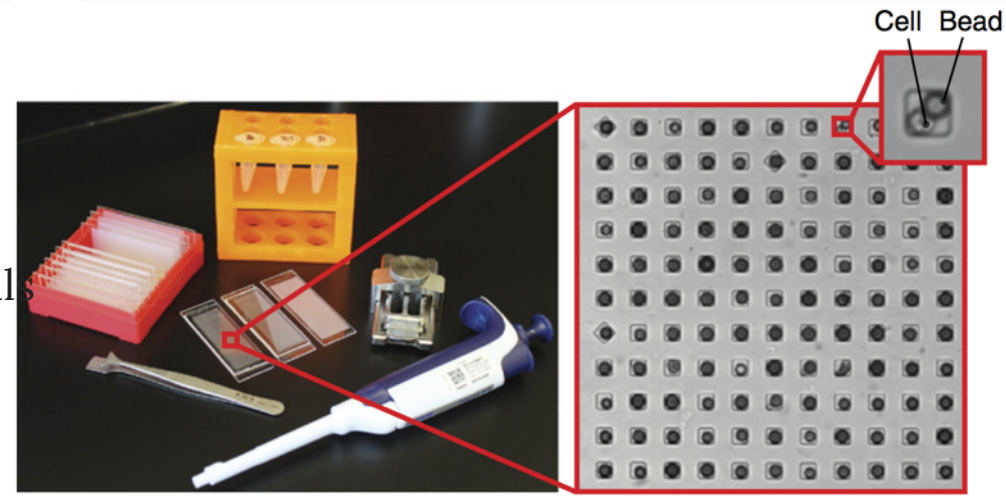
Mouse embryonic development

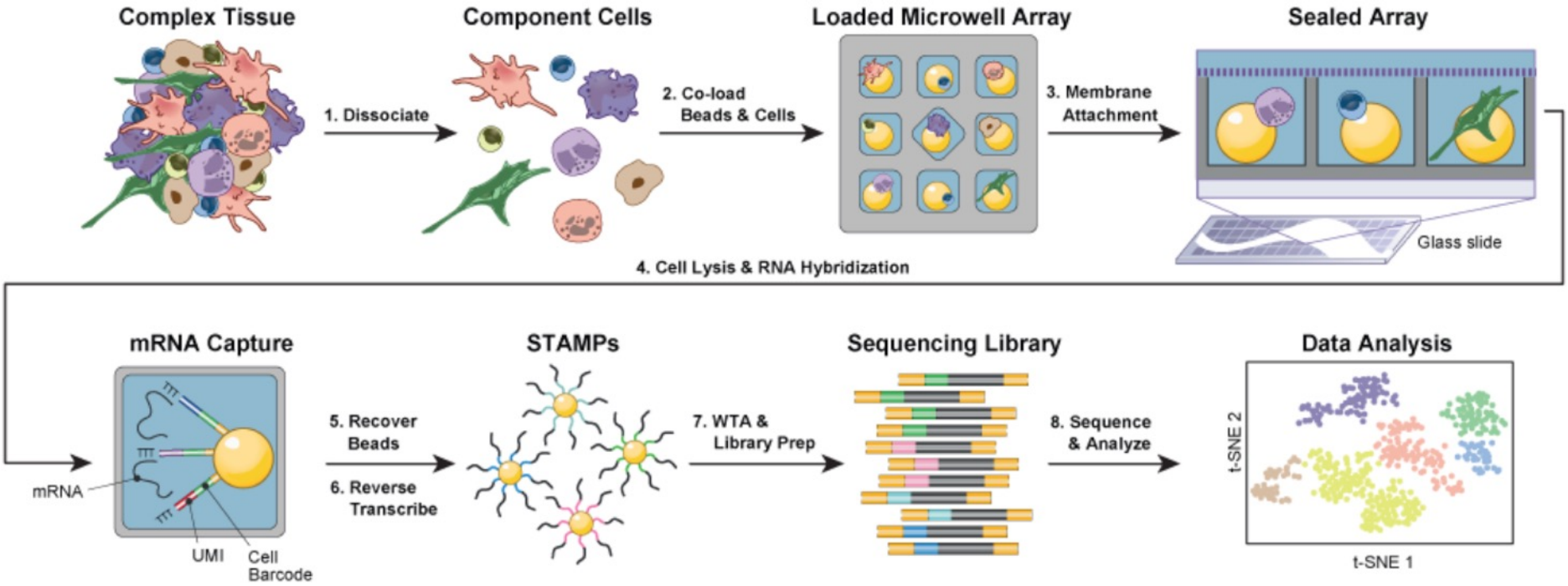


2,058,652 single-cell transcriptomes

Seq-Well

PDMS array of ~86,000 subnanoliter wells
Sized to fit 1 bead per well
Drop-Seq style barcoded beads
Sealed chamber for each cell

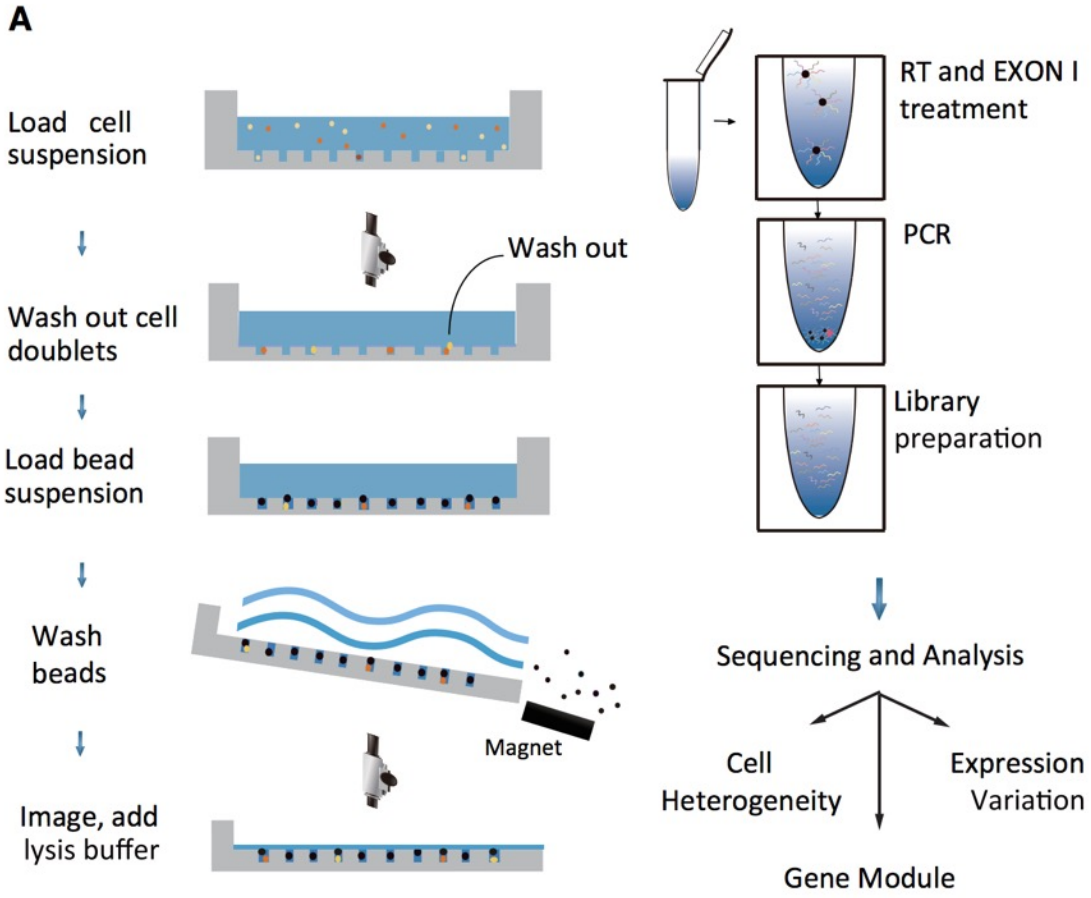


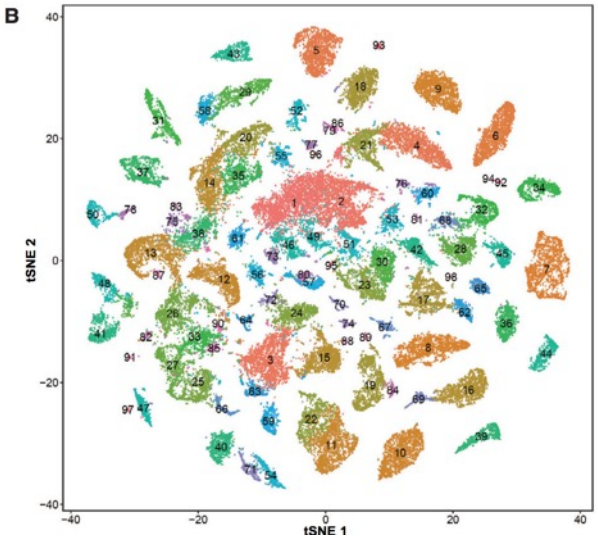


Gierahn et al. *Nature Methods* **14**, 395–398 (2017)

Mapping the Mouse Cell Atlas by Microwell-Seq

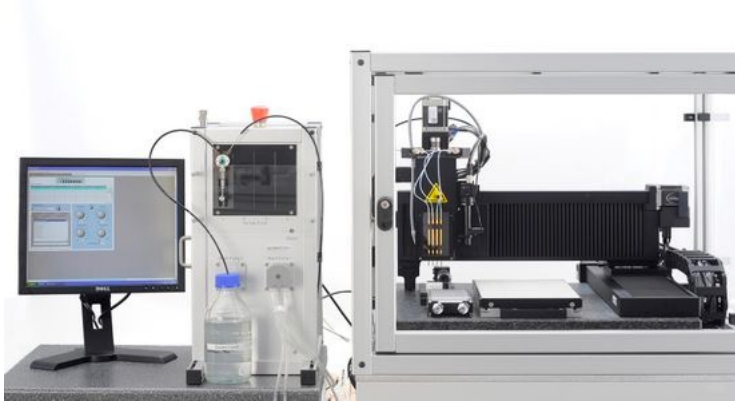
Xiaoping Han,[1,12,13,*] Renying Wang,[1,12,13] Yincong Zhou,[2,12,13] Lijiang Fei,[1,12,13] Huiyu Sun,[1,12,13] Shujing Lai,[1,12,13] Assieh Saadatpour,[11] Ziming Zhou,[1,12] Haide Chen,[1,12] Fang Ye,[1,12] Daosheng Huang,[1] Yang Xu,[1] Wentao Huang,[1] Mengmeng Jiang,[1,12] Xinyi Jiang,[1,12] Jie Mao,[3] Yao Chen,[4] Chenyu Lu,[5] Jin Xie,[6] Qun Fang,[7] Yibin Wang,[8] Rui Yue,[8] Tiefeng Li,[3] He Huang,[9,12] Stuart H. Orkin,[10] Guo-Cheng Yuan,[11] Ming Chen,[2,12] and Guoji Guo[1,9,12,14,*]

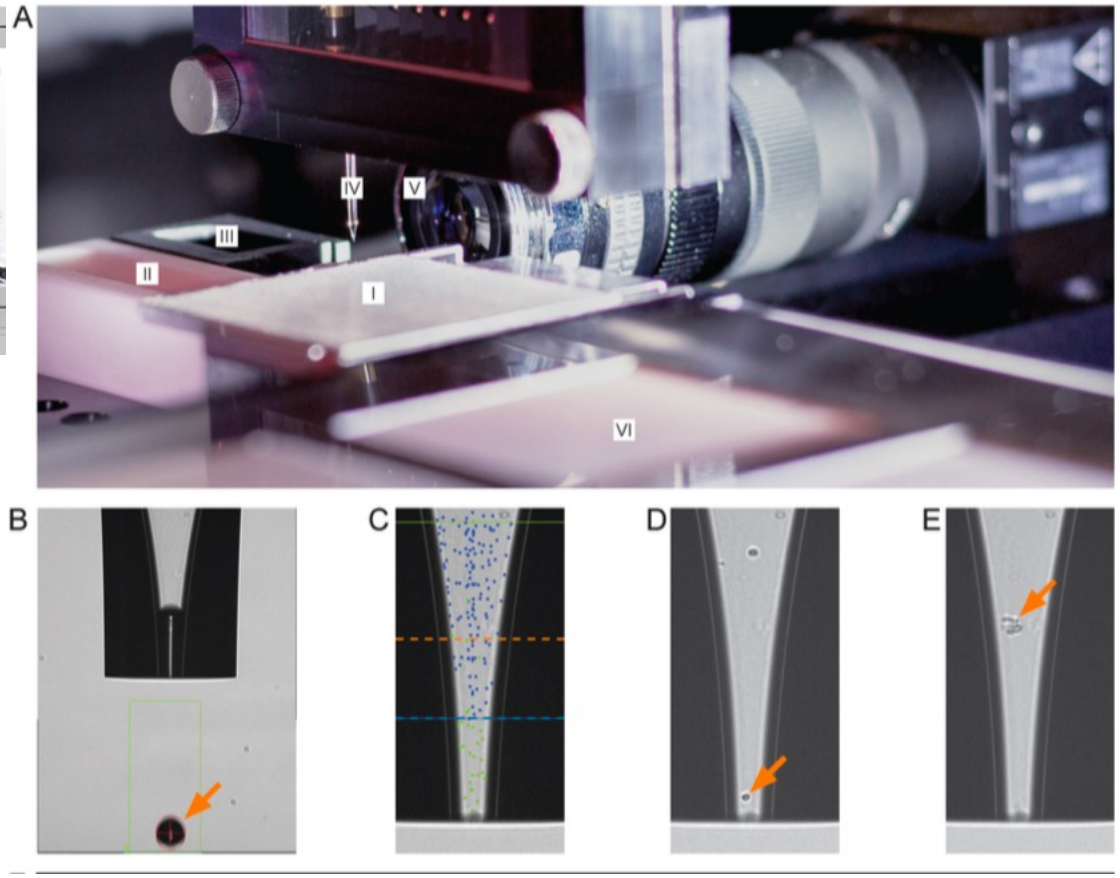




400,000 single cells from major mouse organs

Array-based formats





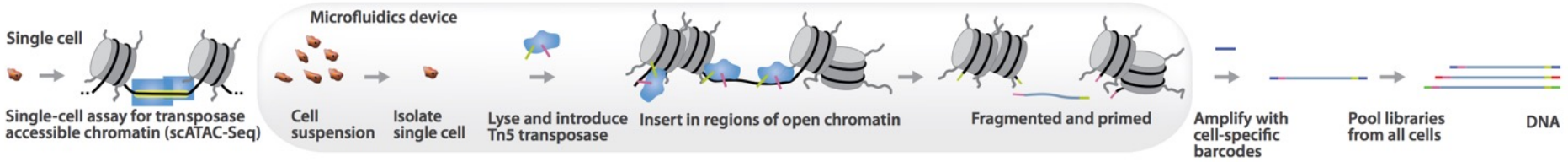
Cell / reagent arrayers

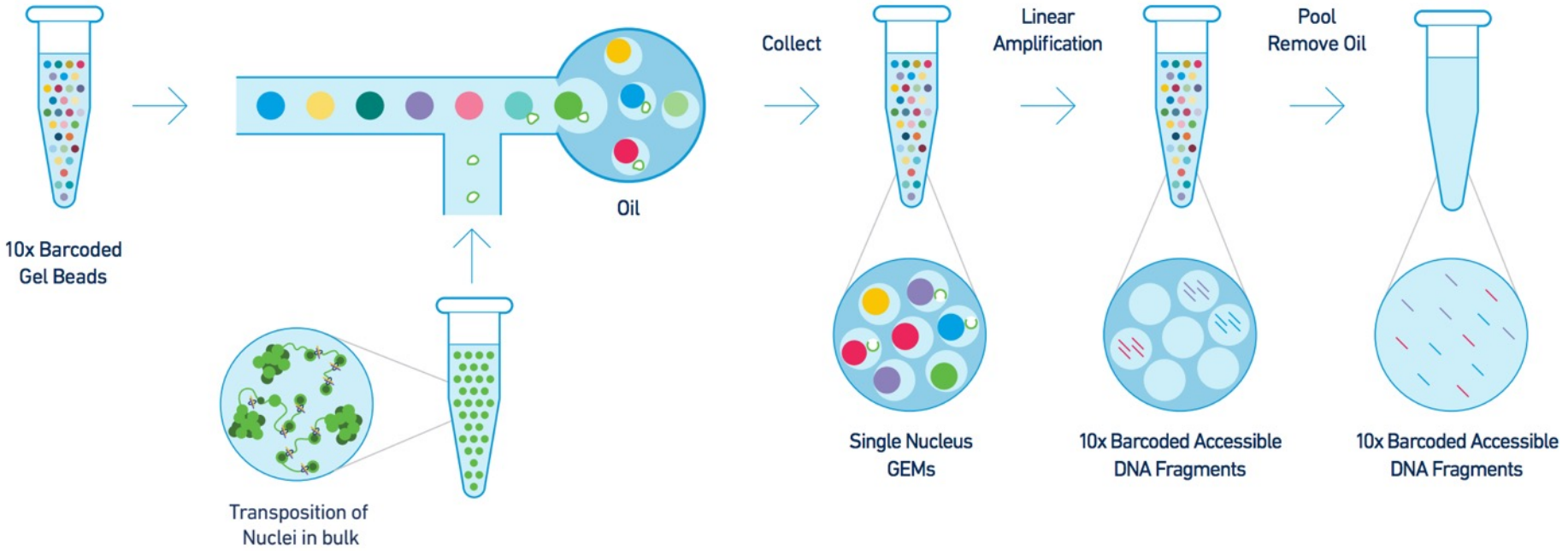
Eg. Scienion sciFlexarray
Scienion cellenONE

Custom workflows
Imaging-based sorting / rejection

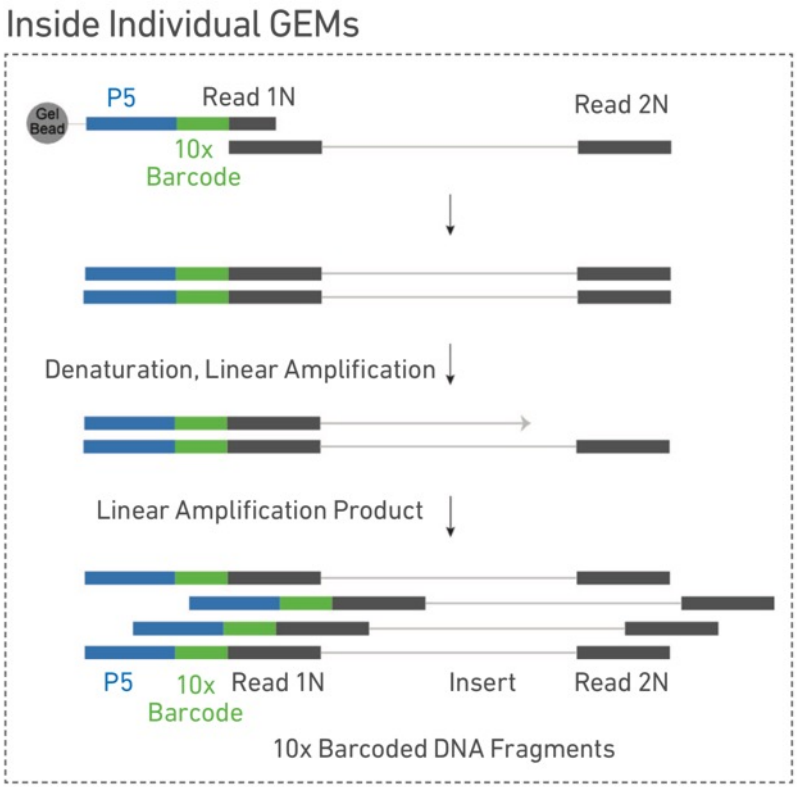

dna
The Other
Nucleic Acid

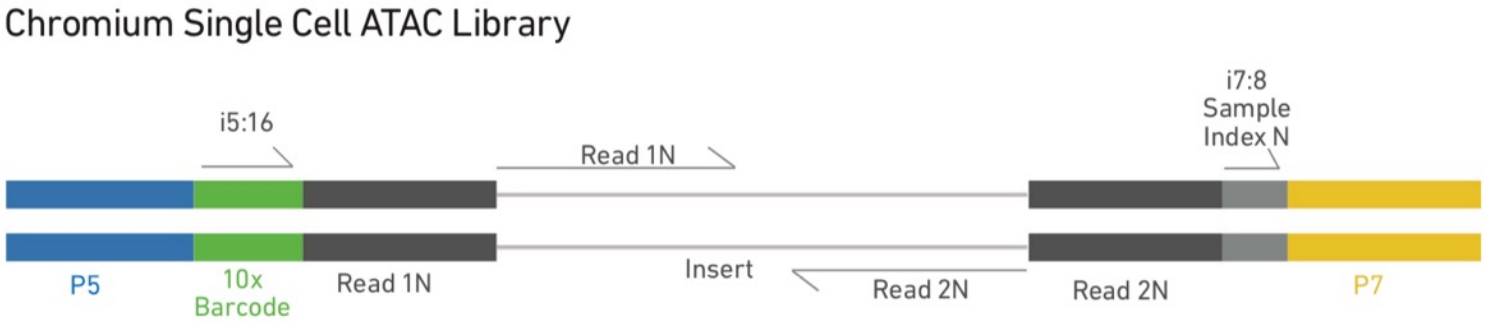
10X Genomics Single Cell ATAC

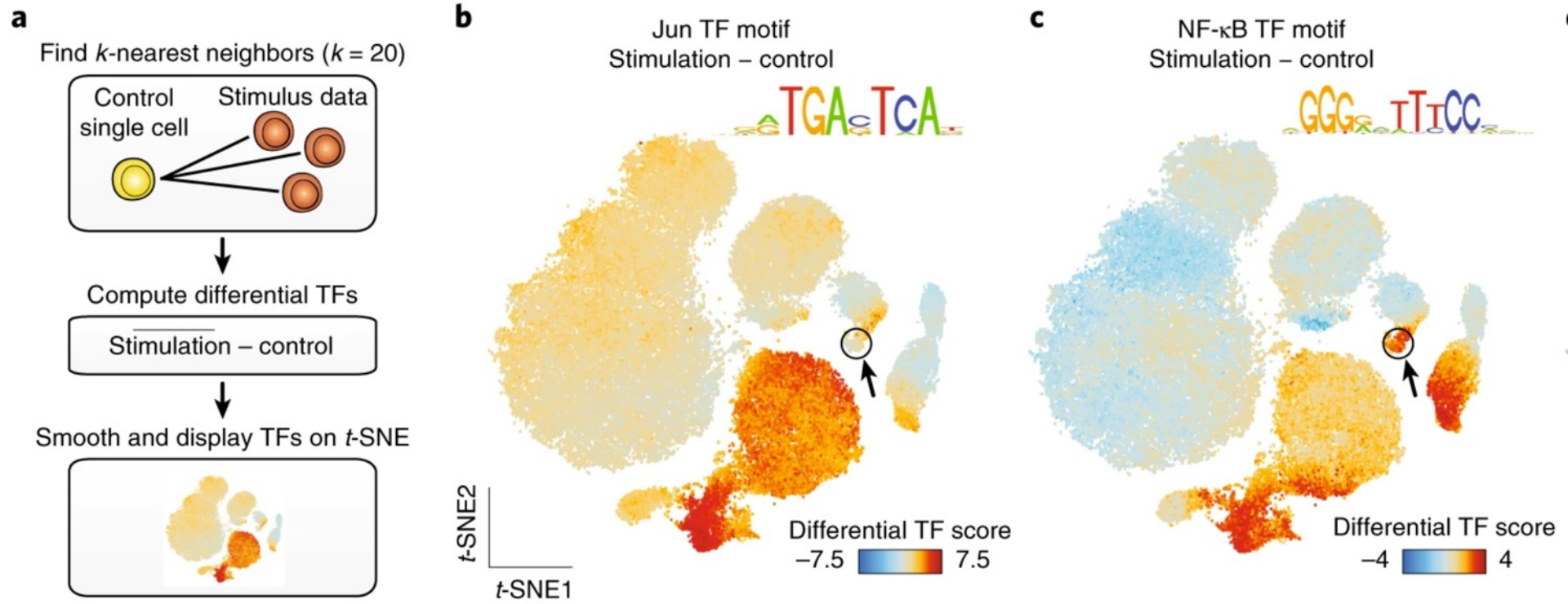




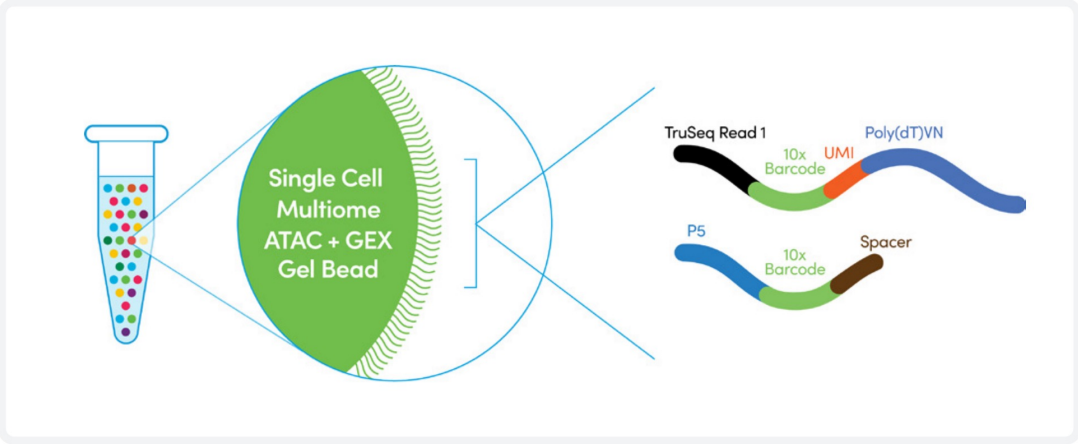
10X Genomics Single Cell ATAC

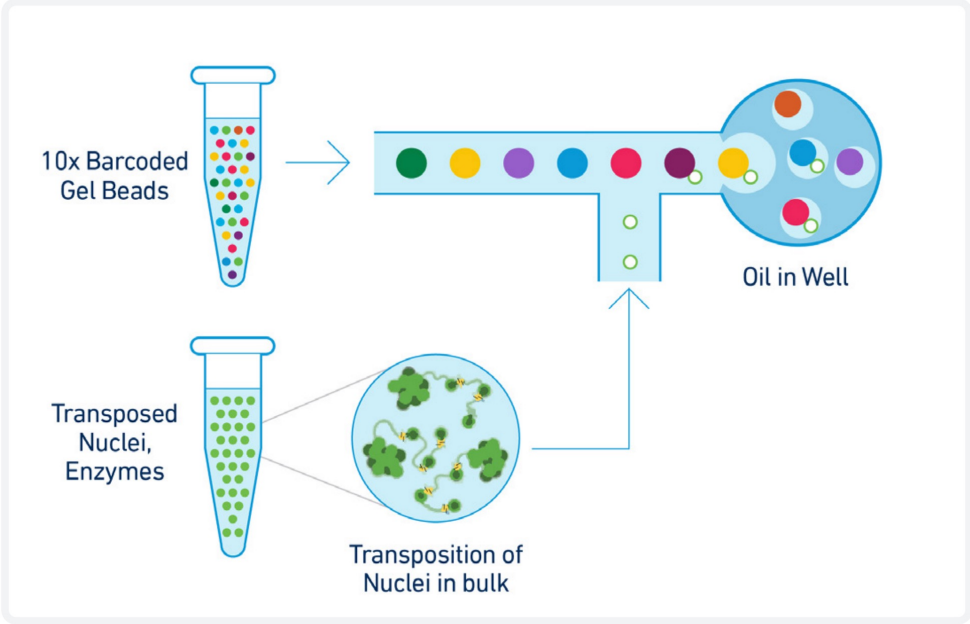




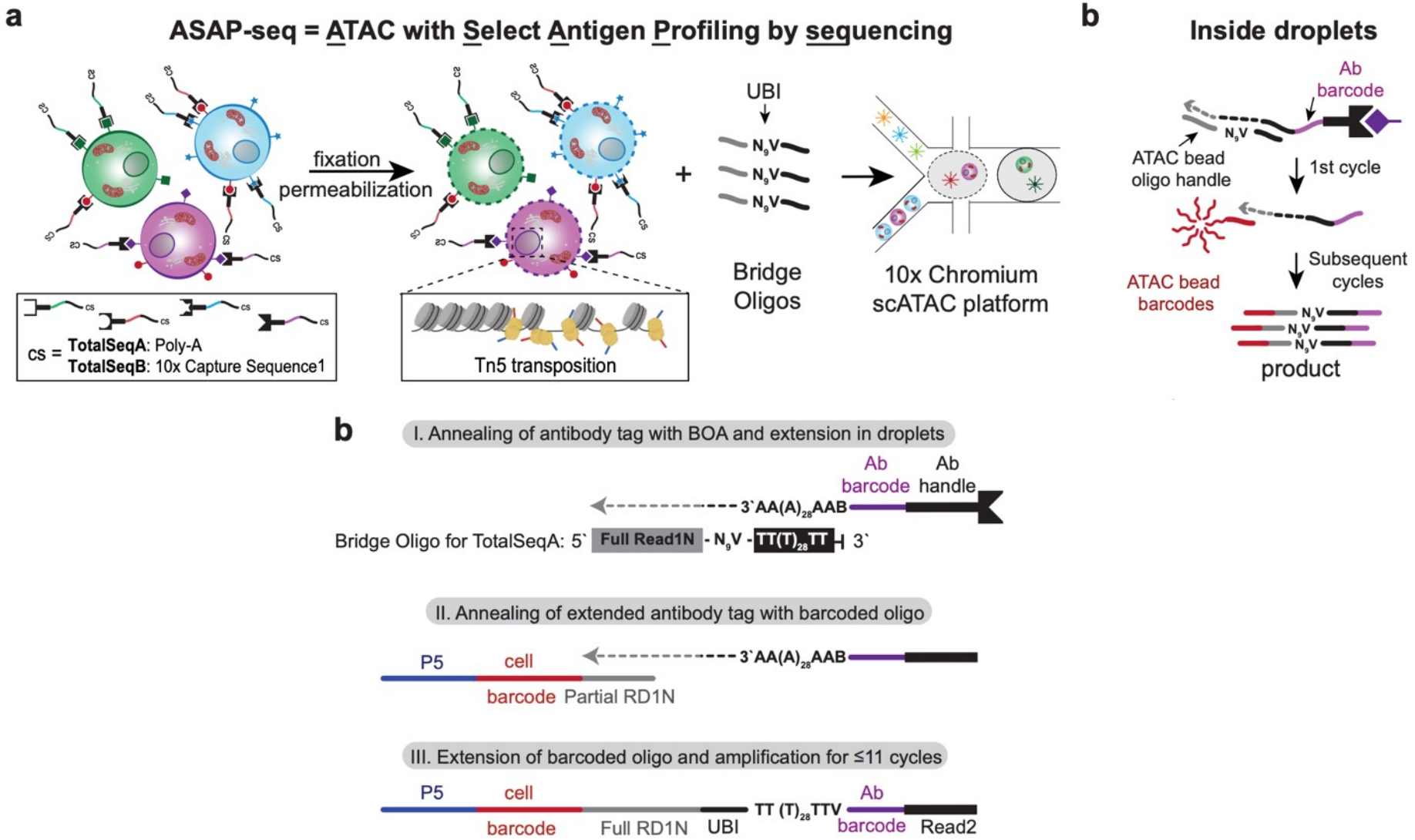

a
Find k-nearest neighbors (k = 20)
Control single cell
Stimulus data
Compute differential TFs
Stimulation – control
Smooth and display TFs on t-SNE
b
Jun TF motif
Stimulation – control
t-SNE2
t-SNE1
Differential TF score
–7.5
7.5
c
NF-κB TF motif
Stimulation – control
Differential TF score
–4
4

10X Genomics Multiome



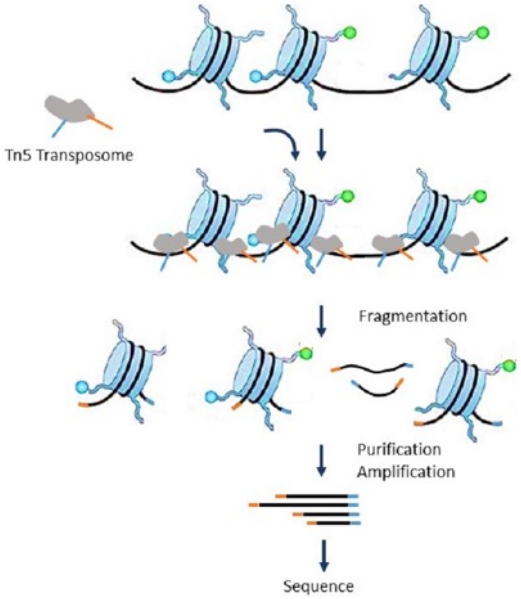


ASAP-seq



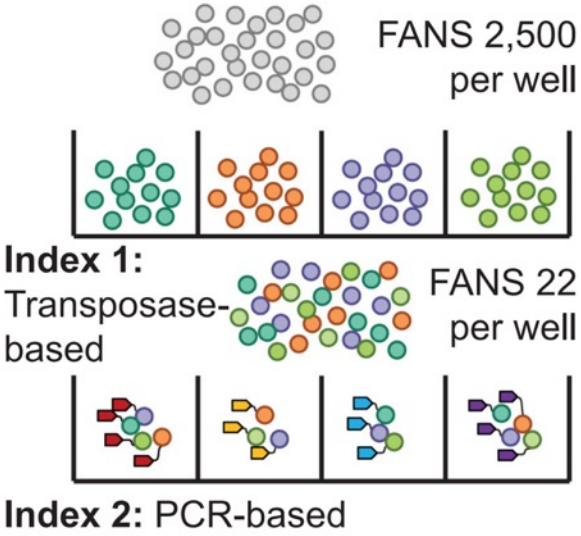
Mimitou et al. bioRxiv (2020) t doi: https://doi.org/10.1101/2020.09.08.286914

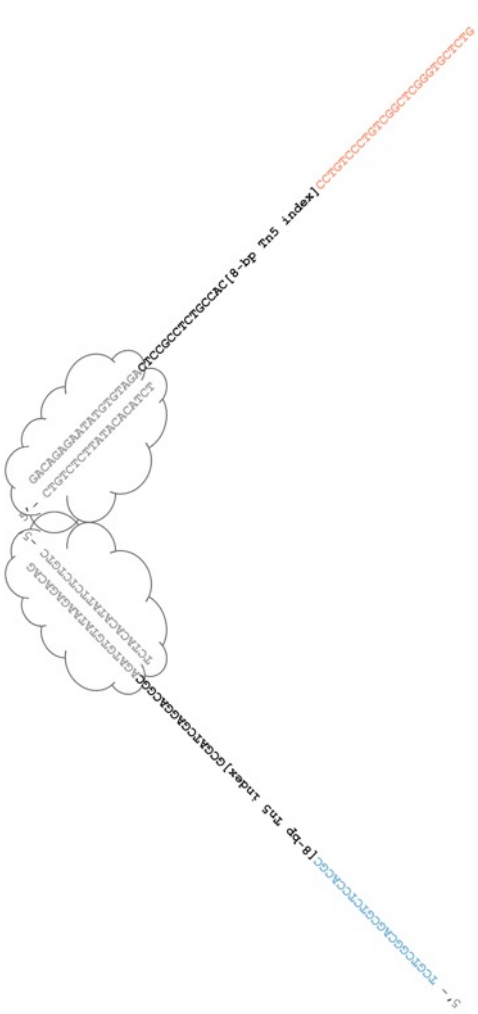
sci-ATAC

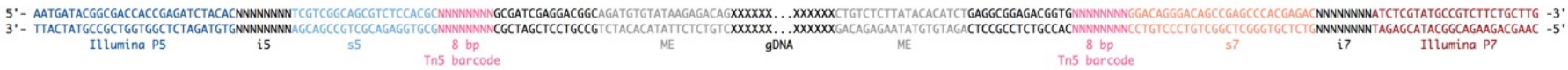


Round 1:
Internally Barcoded
Tn5 transposomes

Round 2:
Barcoded PCR primers

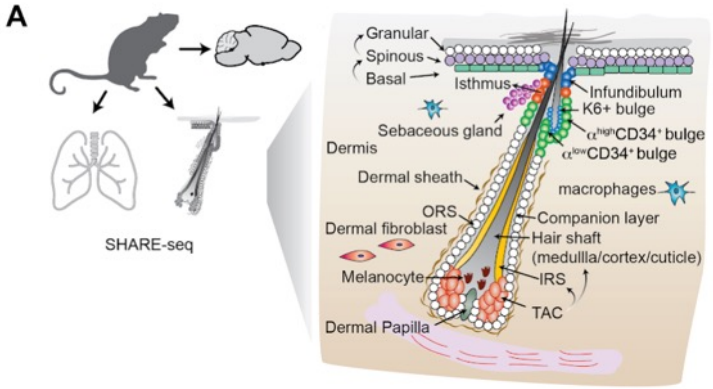


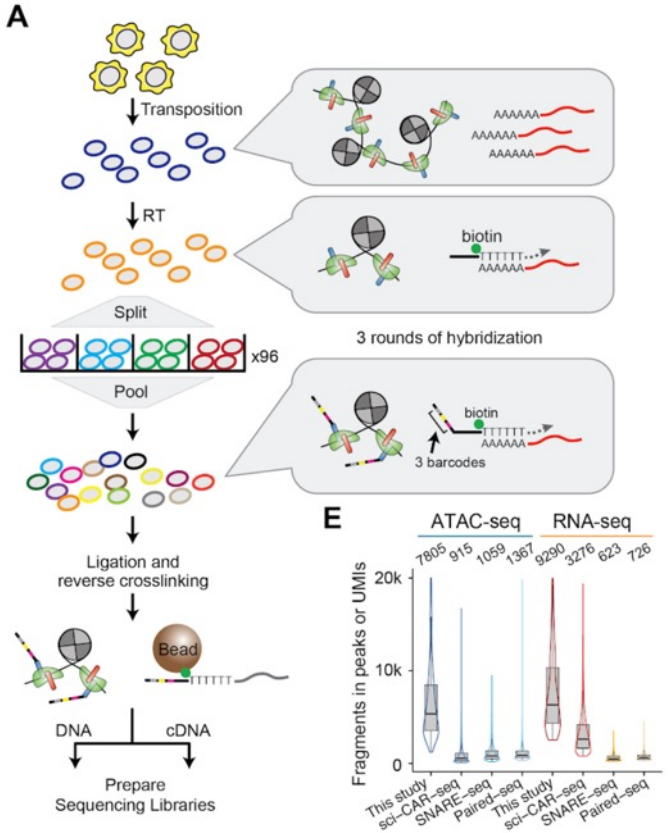


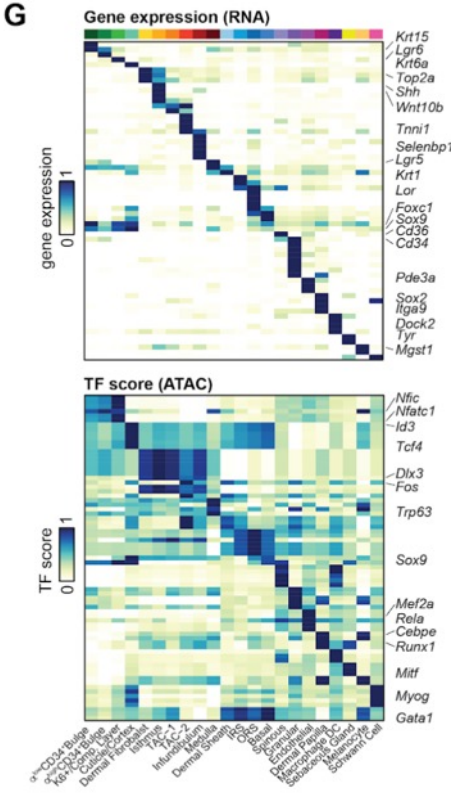


SHARE-Seq

- Same-cell scRNA/ATAC
- Combinatorial split-pool barcoding of adapters

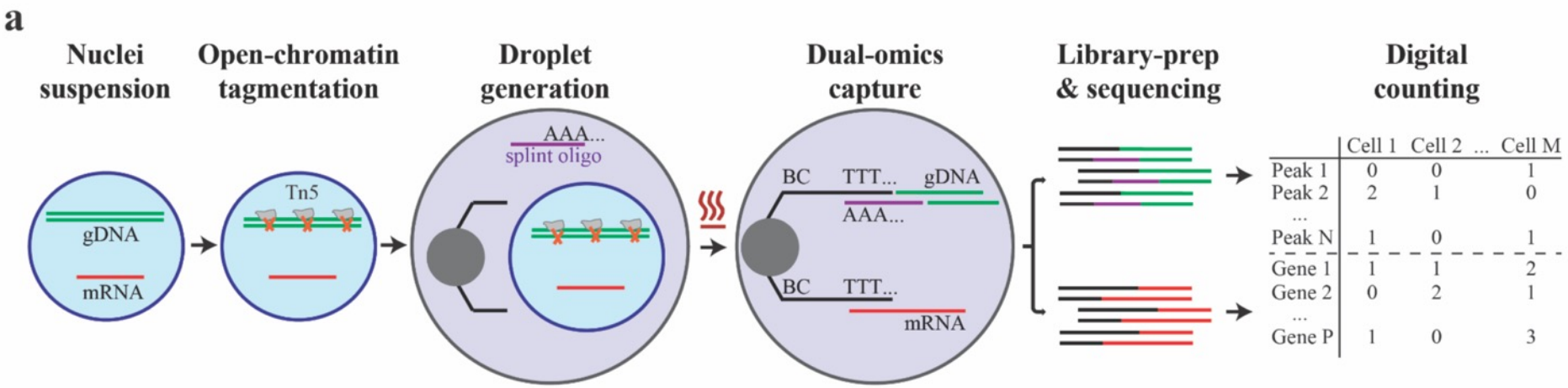


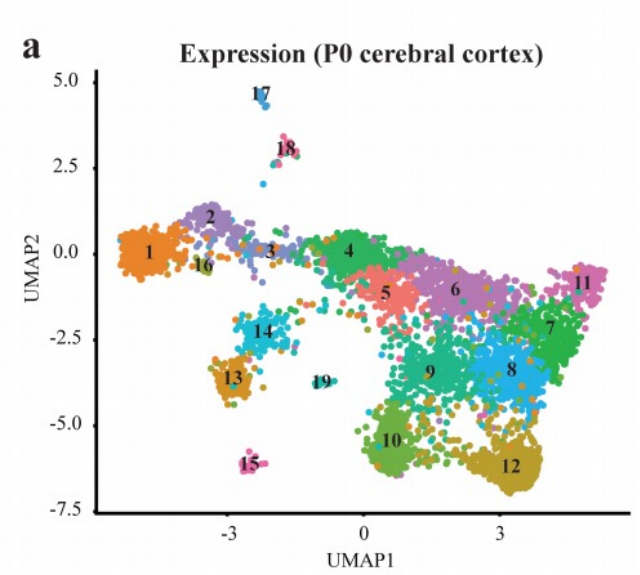


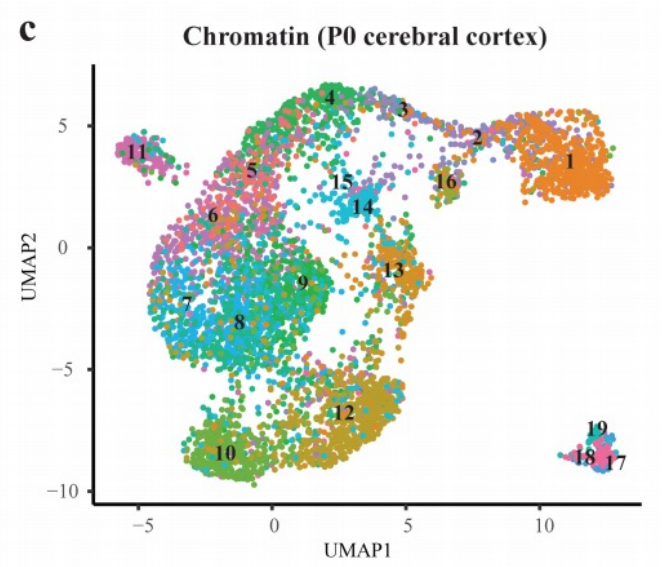


Ma et al. Cell. 2020 Oct 20;S0092-8674(20)31253-8. doi:10.1016/j.cell.2020.09.056.

SNARE-seq



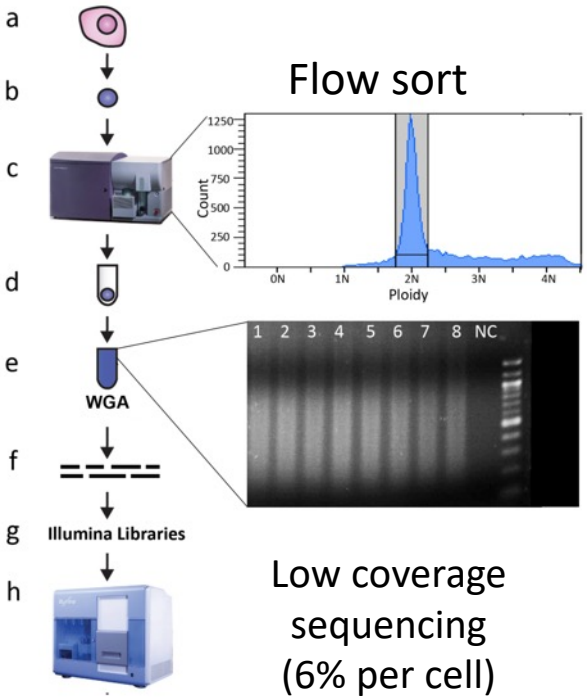




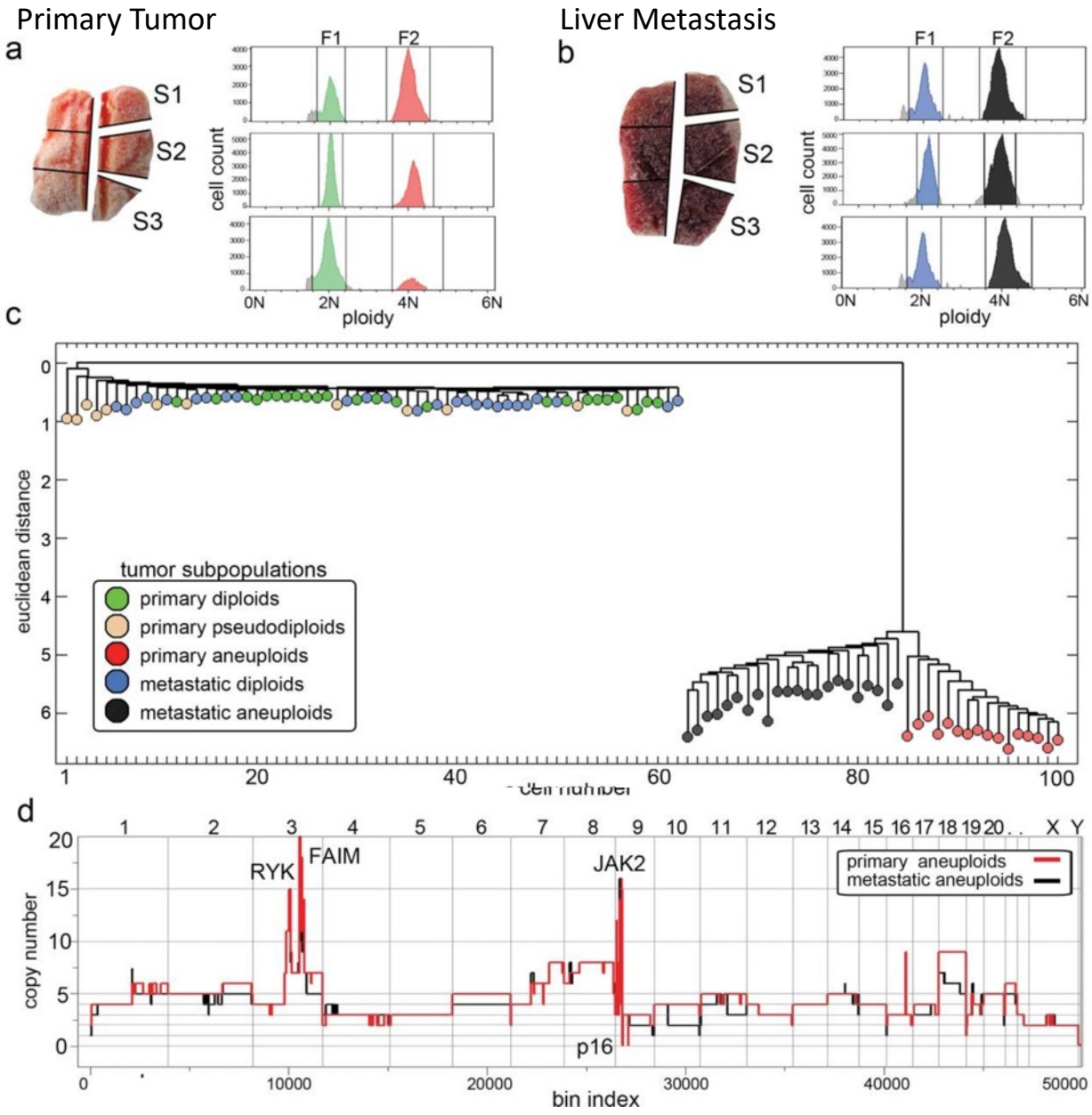
Chen et al. *bioRxiv* (**2019**) doi:10.1101/692608

Single cell CNV

Nick Navin, Mike Wigler
CSHL

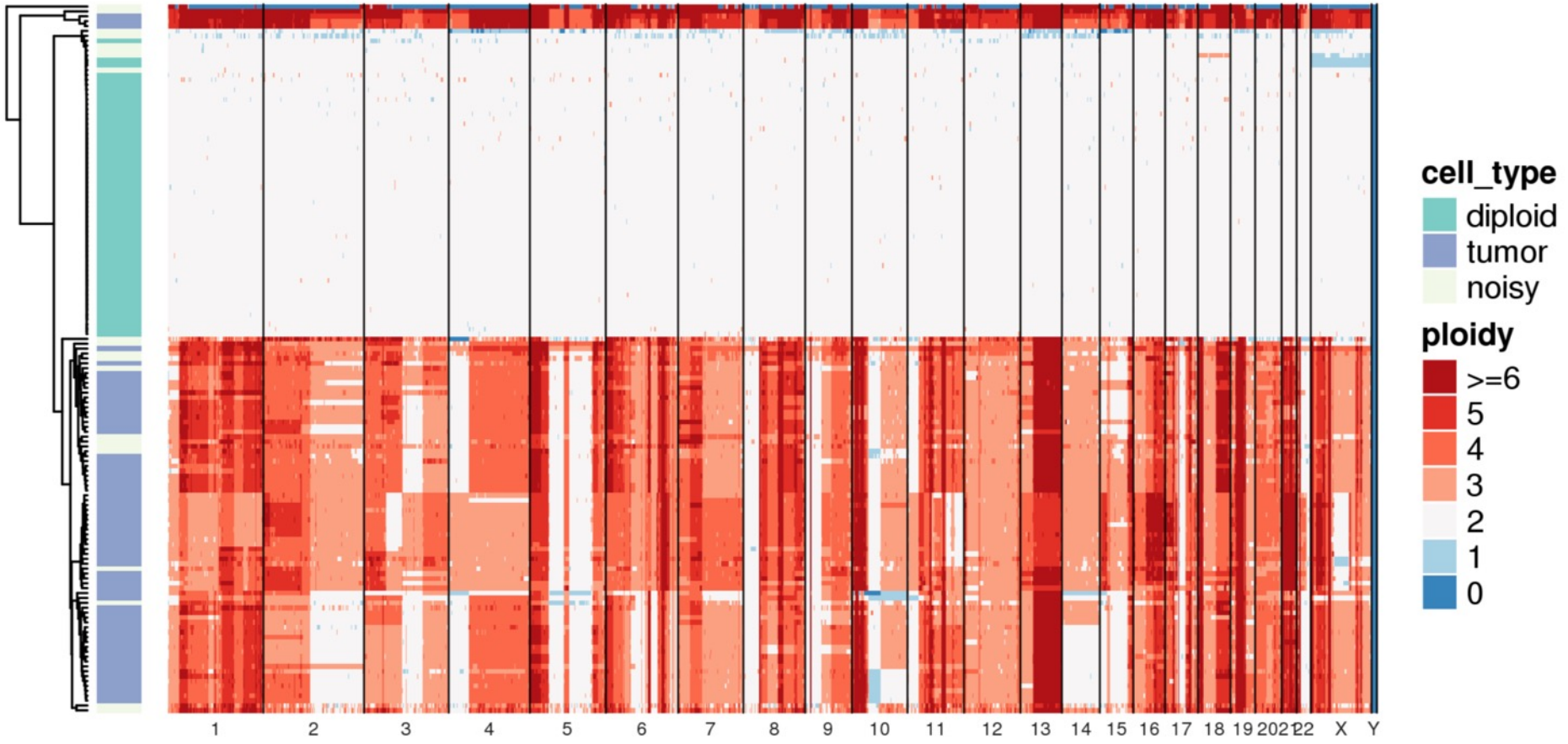


Low coverage sequencing (6% per cell)

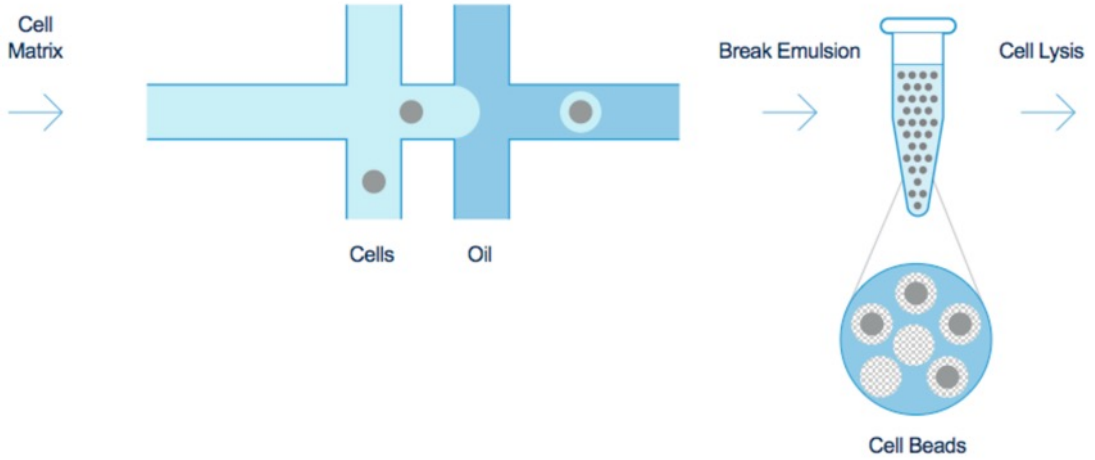


Navin et al. *Nature* (**2011**) Apr 7; 472(7341): 90–94.

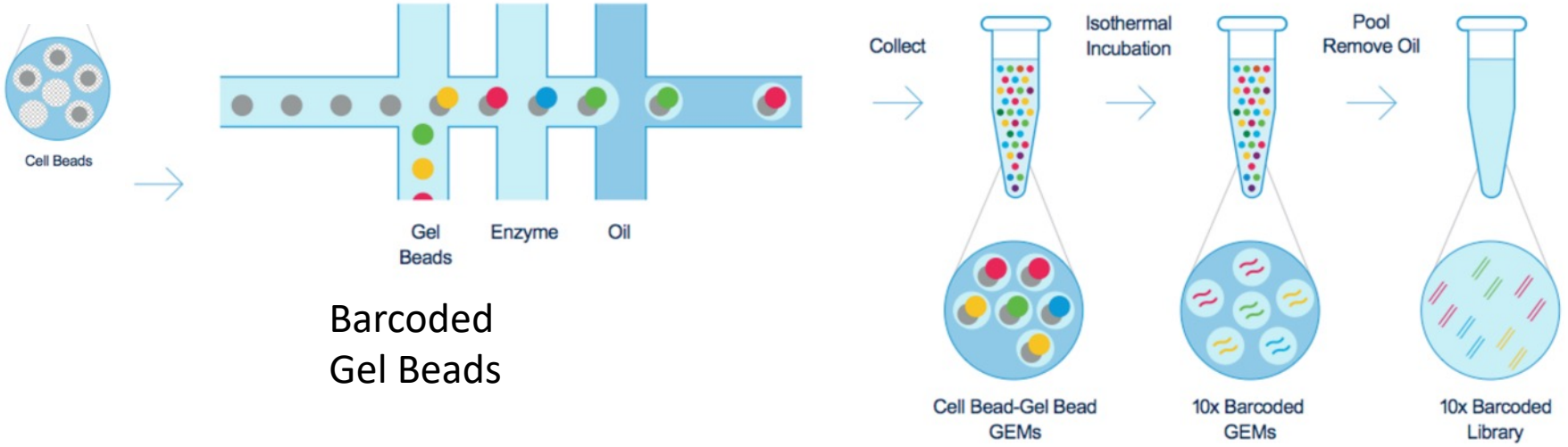
Droplet-based Single Cell CNV



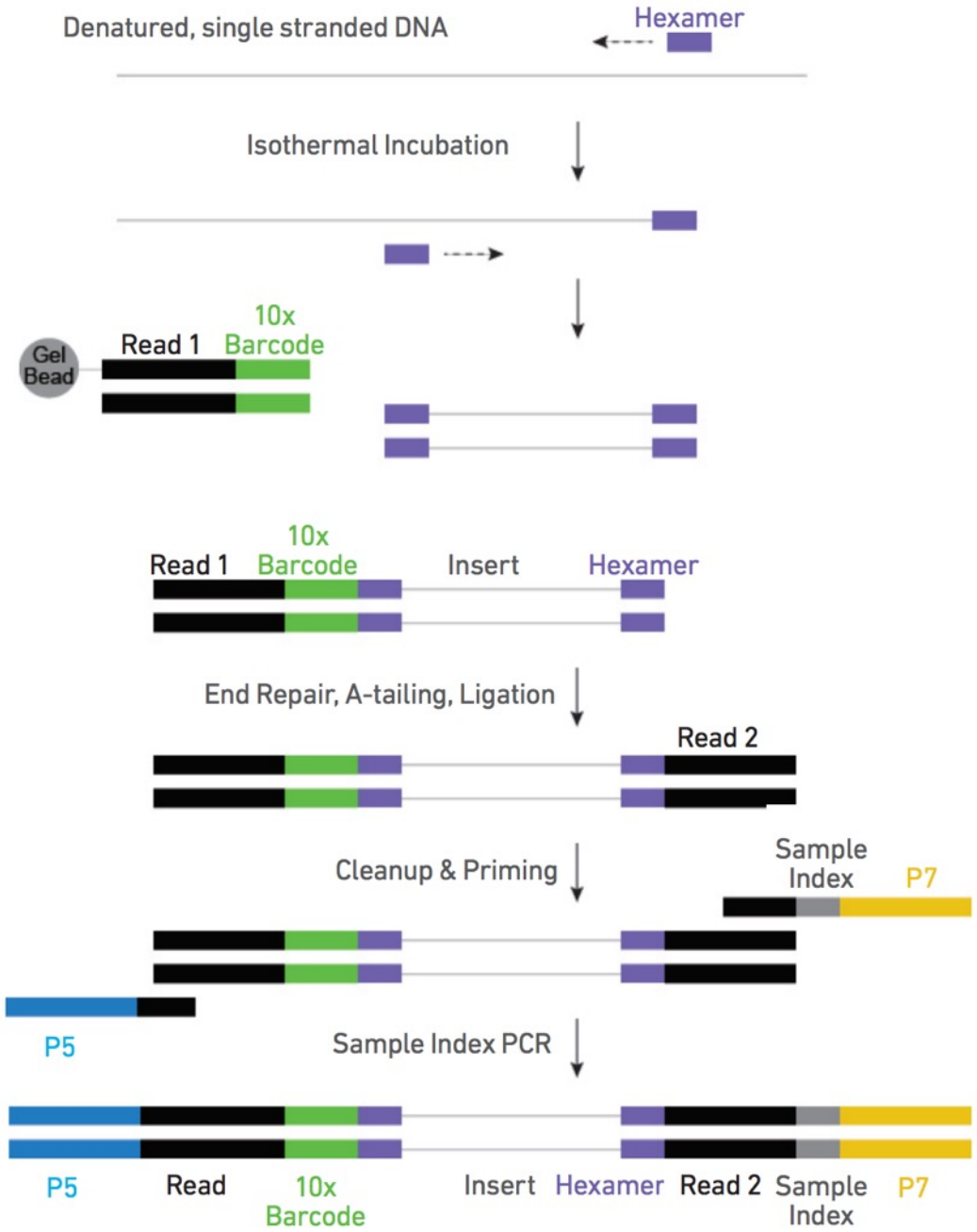
Droplet-based Single Cell CNV



Deproteinized, lysed cell w/DNA trapped in bead



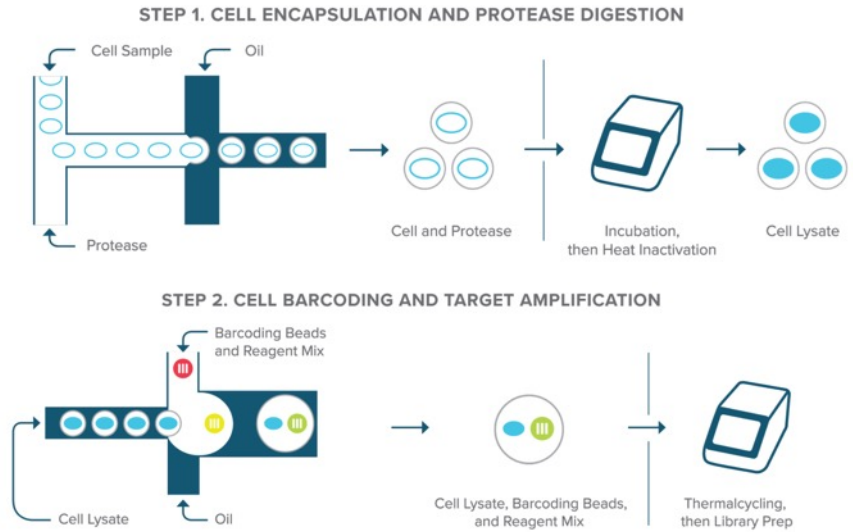
Barcoded
Gel Beads


Denatured, single stranded DNA
Hexamer
Isothermal Incubation
10x
Read 1
Barcode
Gel
Bead
10x
Read 1
Barcode
Insert
Hexamer
End Repair, A-tailing, Ligation
Read 2
Cleanup & Priming
Sample
Index
P7
P5
Sample Index PCR
P5
Read
10x
Barcode
Insert
Hexamer
Read 2
Sample
Index
P7

Mission Bio Tapestri

DNA-focused microfluidic platform

For SNV & CNV



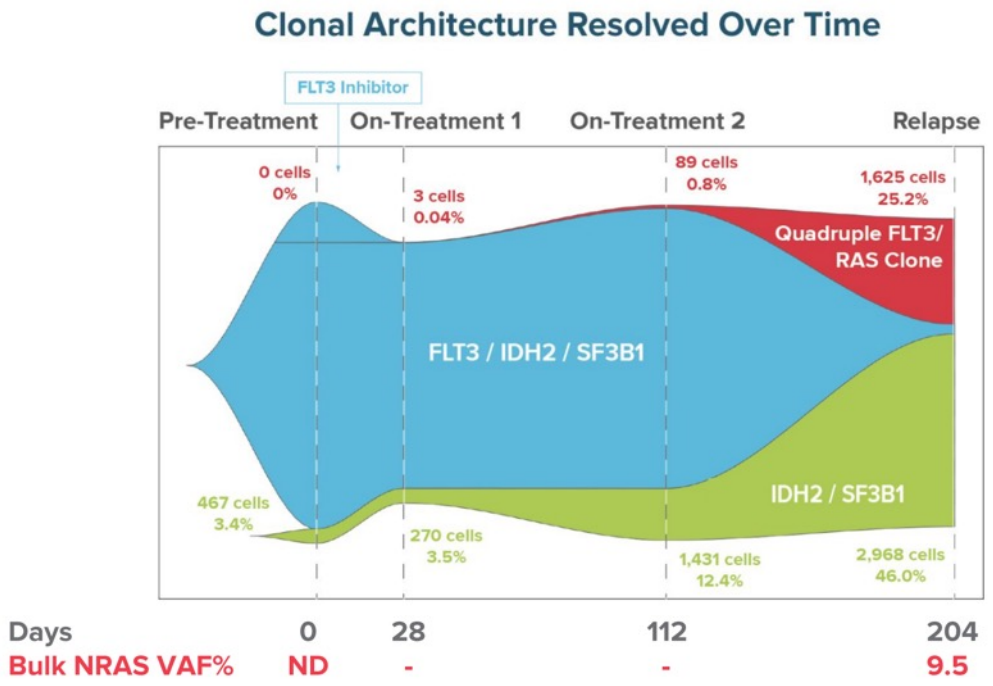
Mission Bio Tapestri

59 GENES - TUMOR HOTSPOT PANEL

ABL1	CSF1R	FGFR1	IDH2	MLH1	RB1
AKT1	CTNNB1	FGFR2	JAK1	MPL	RET
ALK	DDR2	FGFR3	JAK2	MTOR	SMAD4
APC	EGFR	FLT3	JAK3	NOTCH1	SMARCB1
AR	ERBB2	GNA11	KDR	NRAS	SMO
ATM	ERBB3	GNAQ	KIT	PDGFRA	SRC
BRAF	ERBB4	GNAS	KRAS	PIK3CA	STK11
CDH1	ESR1	HNF1A	MAP2K1	PTEN	TP53
CDK4	EZH2	HRAS	MAP2K2	PTPN11	VHL
CDKN2A	FBXW7	IDH1	MET	RAF1	

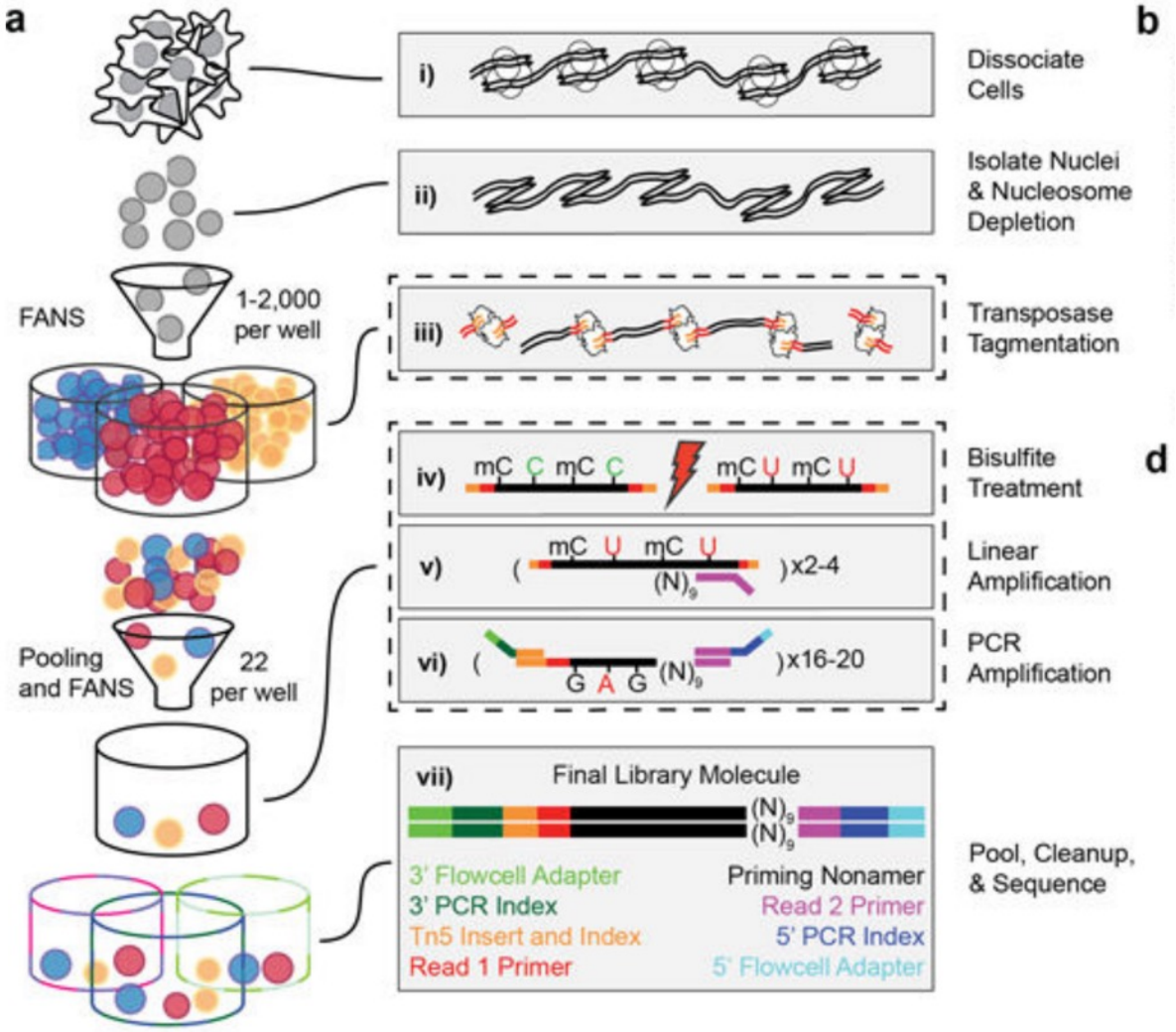
45-GENE MYELOID PANEL

ASXL1	ERG	KDM6A	NRAS	SMC1A
ATM	ETV6	KIT	PHF6	SMC3
BCOR	EZH2	KMT2A	PPM1D	STAG2
BRAF	FLT3	KRAS	PTEN	STAT3
CALR	GATA2	MPL	PTPN11	TET2
CBL	GNAS	MYC	RAD21	TP53
CHEK2	IDH1	MYD88	RUNX1	U2AF1L5
CSF3R	IDH2	NF1	SETBP1	WT1
DNMT3A	JAK2	NPM1	SF3B1	ZRSR2



Other Omics

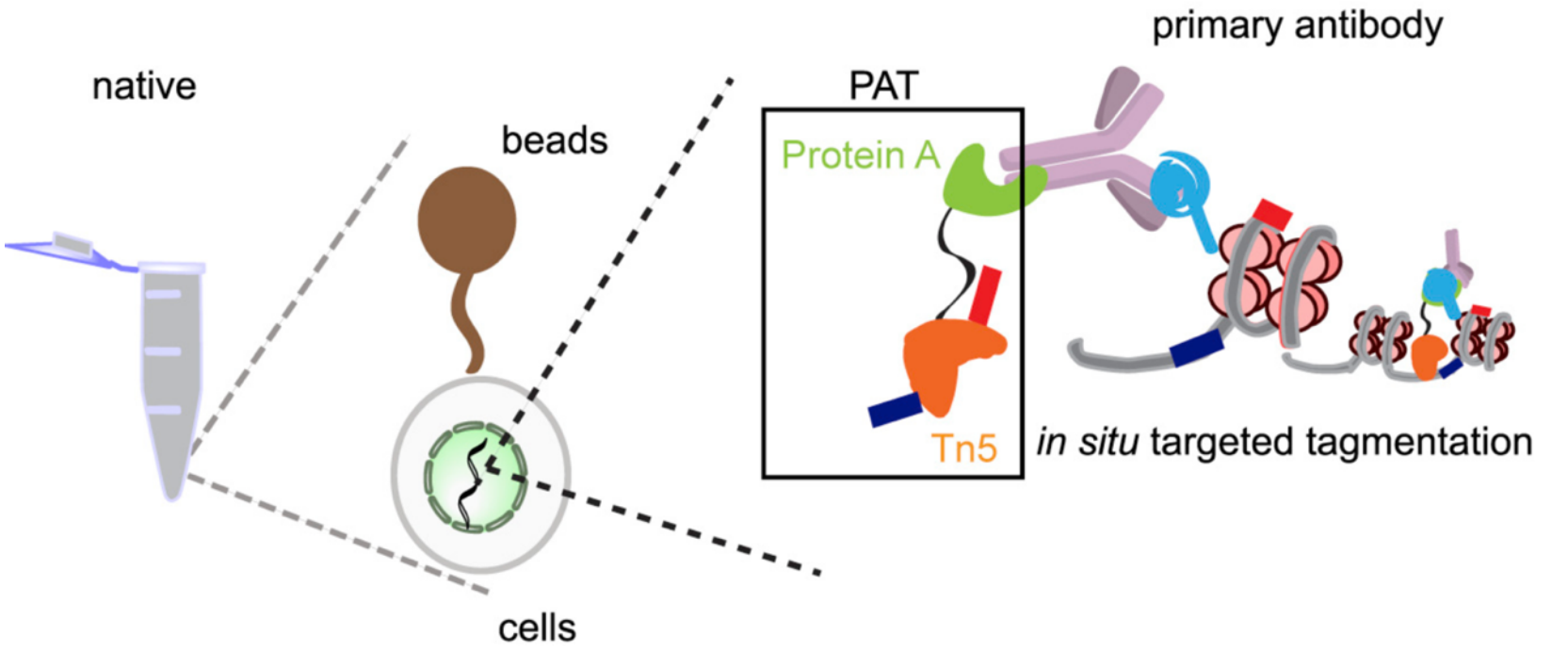
sci-MET



Mulqueen et al. *Nat Biotechnol*. **2018** Jun;36(5):428-431

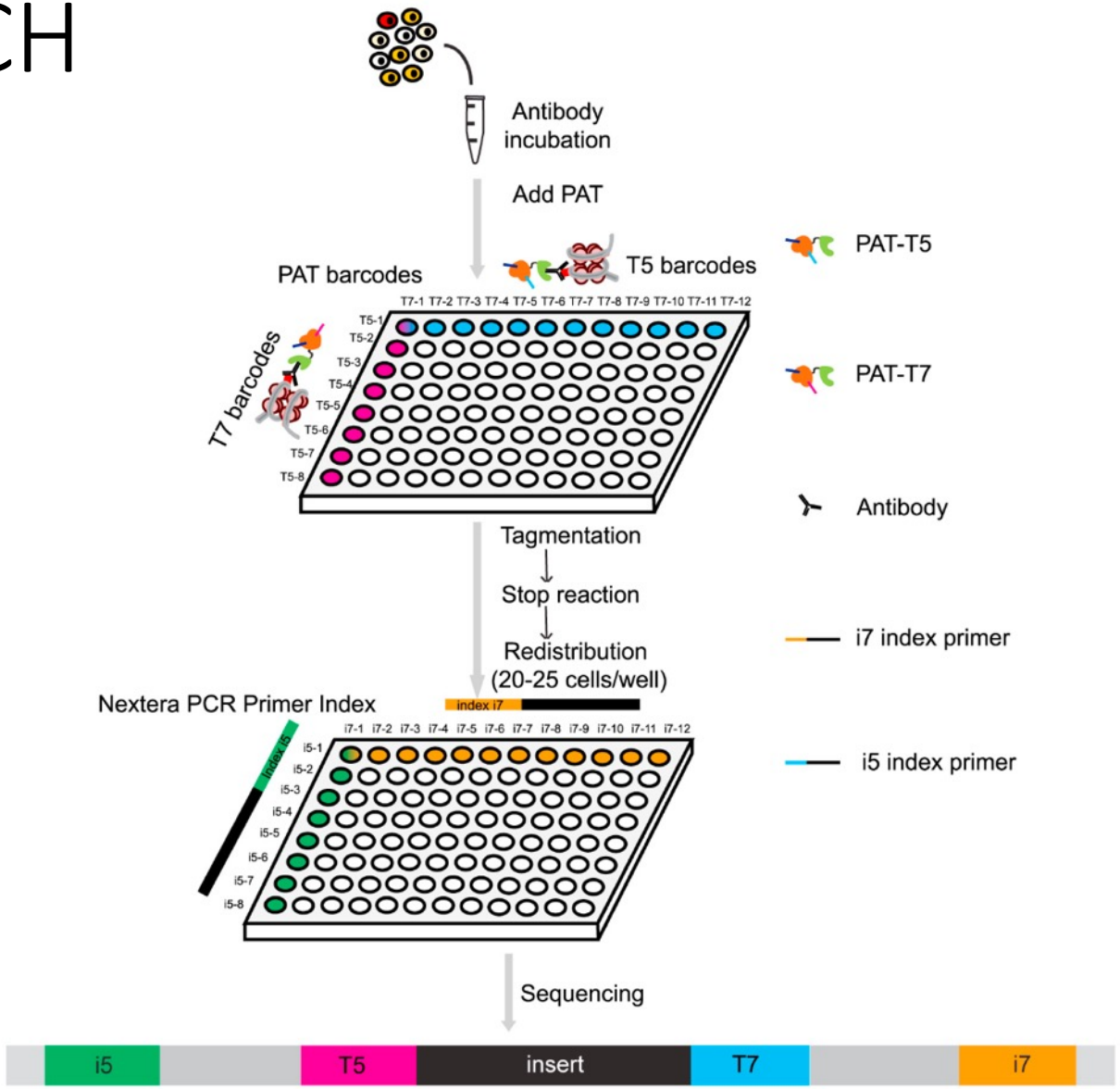
CoBATCH

- Transcription factor binding sites in single cells

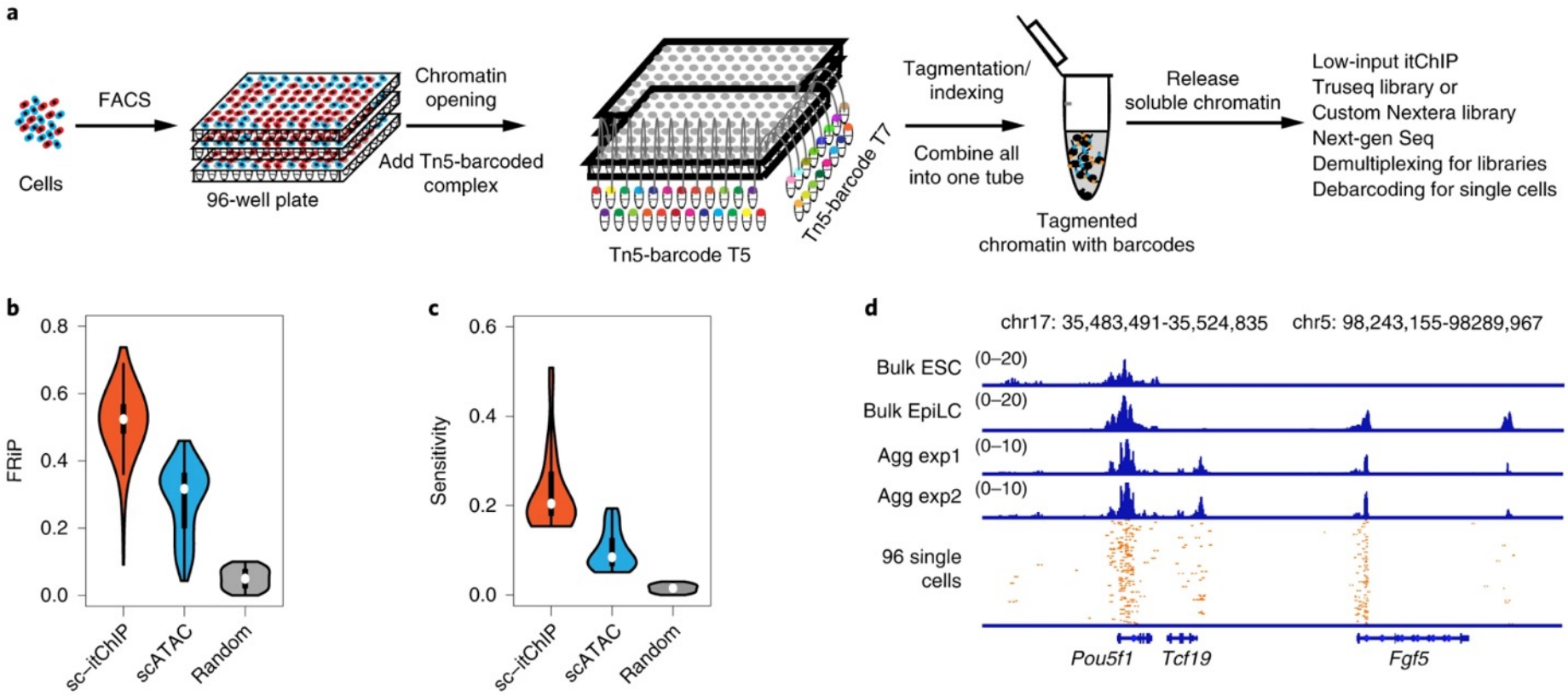


Wang et al *Mol Cell* (2019) **76**:1, p206-216.e7

CoBATCH


Antibody incubation
Add PAT
PAT barcodes
T5 barcodes
T7 barcodes
T7-1 T7-2 T7-3 T7-4 T7-5 T7-6 T7-7 T7-8 T7-9 T7-10 T7-11 T7-12
T5-1
T5-2
T5-3
T5-4
T5-5
T5-6
T5-7
T5-8
Tagmentation
Stop reaction
Redistribution
(20-25 cells/well)
Nextera PCR Primer Index
index i7
index i5
i7-1 i7-2 i7-3 i7-4 i7-5 i7-6 i7-7 i7-8 i7-9 i7-10 i7-11 i7-12
i5-1
i5-2
i5-3
i5-4
i5-5
i5-6
i5-7
i5-8
Sequencing
i5
T5
insert
T7
i7
PAT-T5
PAT-T7
Antibody
i7 index primer
i5 index primer

Single cell itChIP



Ai et al. *Nat Cell Biol* (2019) **21**, p1164–1172

Sequencing Costs

	RNA-seq	ATAC-seq	CNV
Reads per Cell	50-100k	50-100k	750k+
Cells per Experiment	2,000 – 10,000	2,000 – 10,000	1,000-2,000
Sequencing Platform Min.	NextSeq HO	NextSeq HO	NovaSeq S1
Cost per Experiment	~$2,500	~$2,800	$12,000