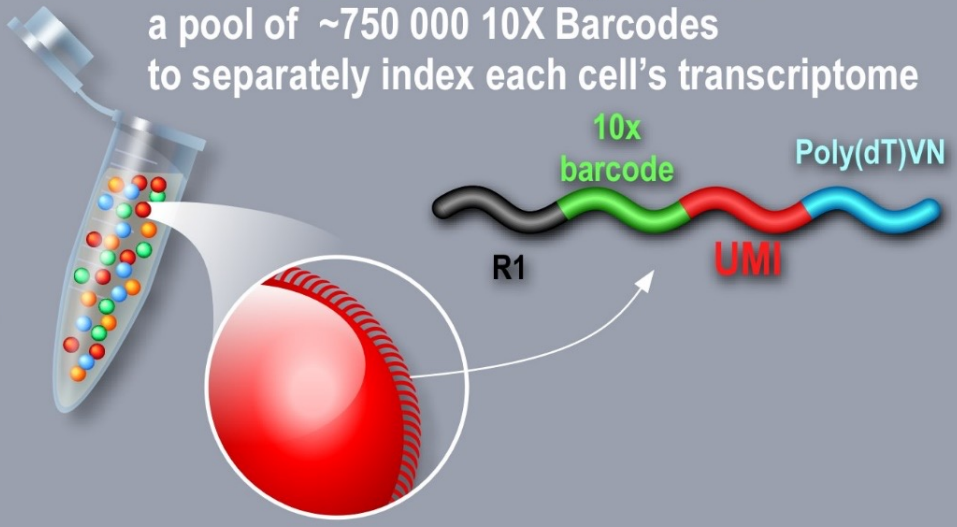
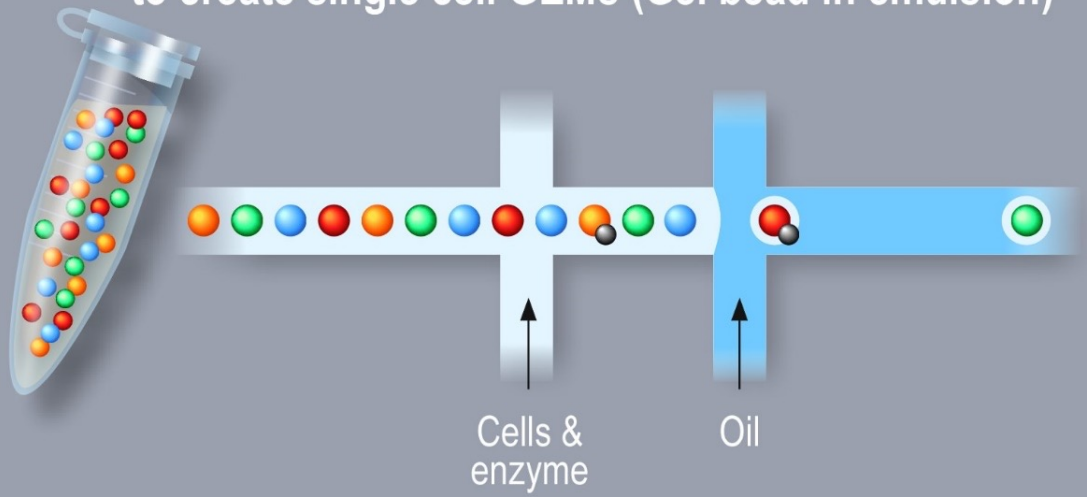
10X Genomics Chromium

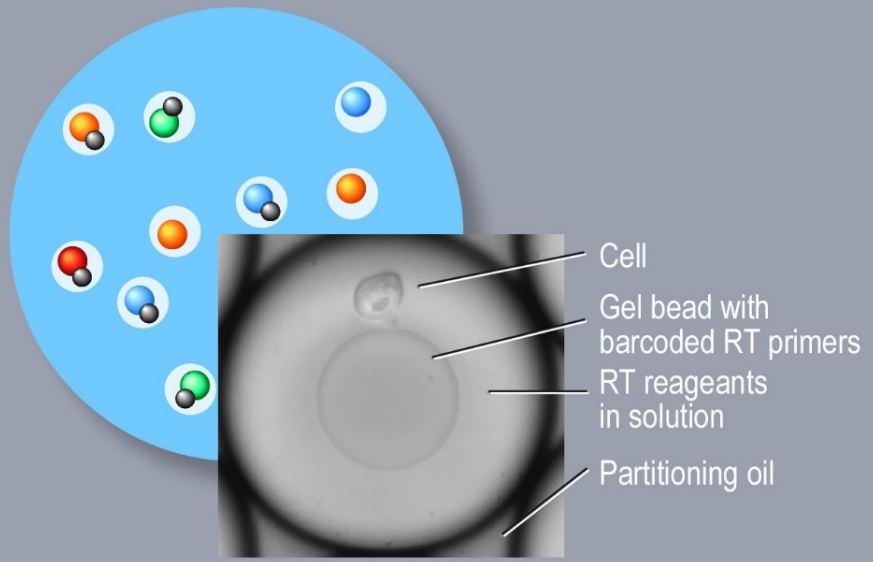
Single-use microfluidics chip

10X GemCode Technology samples
a pool of ~750 000 10X Barcodes
to separately index each cell's transcriptome



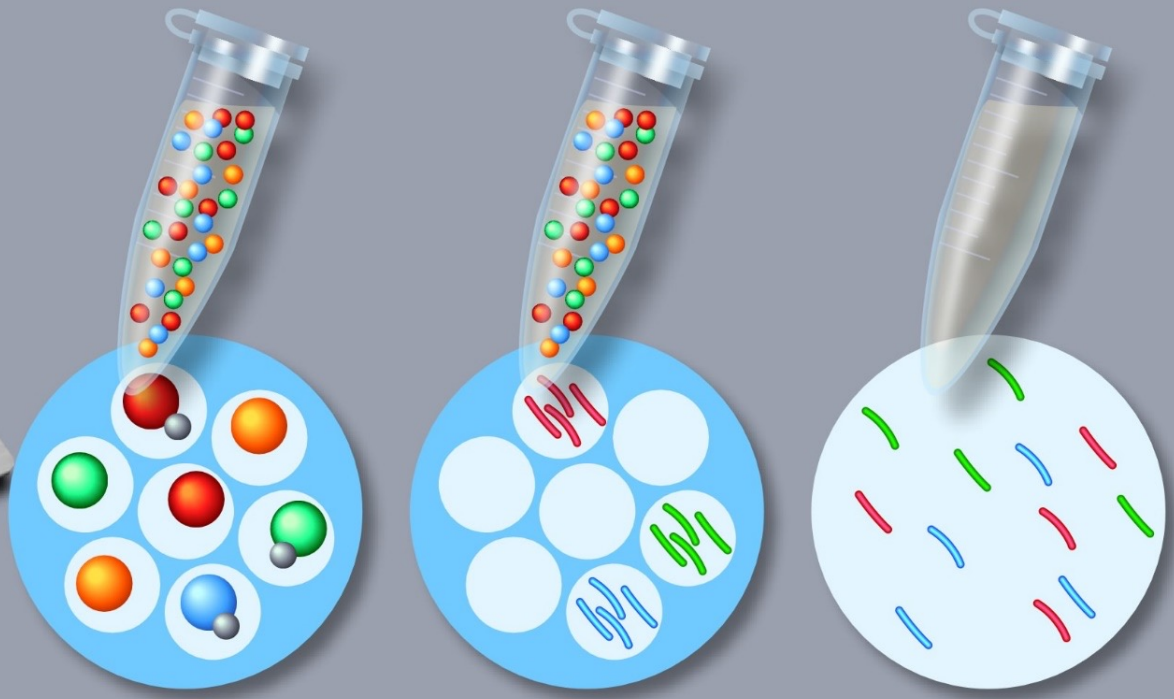
10X Barcoded Gel Beads are mixed with cells, enzyme and oil
to create single cell GEMs (Gel bead in emulsion)

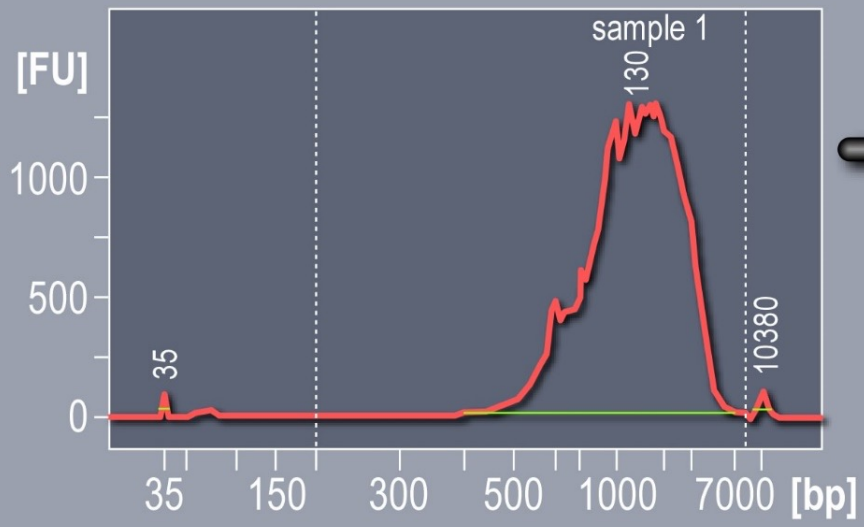




GEMs recovery and cDNA amplification

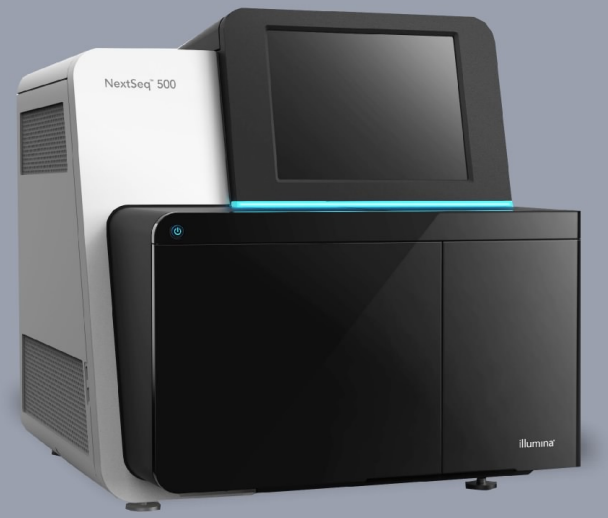
GEM recovery RT
cDNA Amplification
QC and Quantification

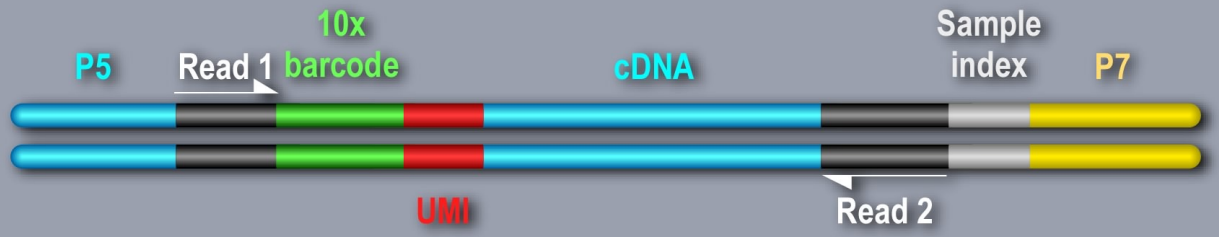




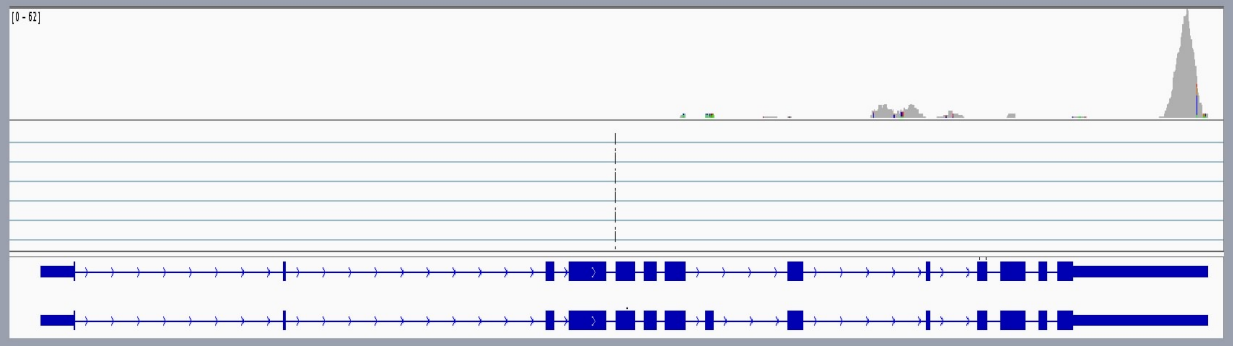
Library construction and Illumina sequencing

Fragment, End Repair and A-tailing
Adaptor Ligation
Sample Index PCR
QC and qPCR quantification





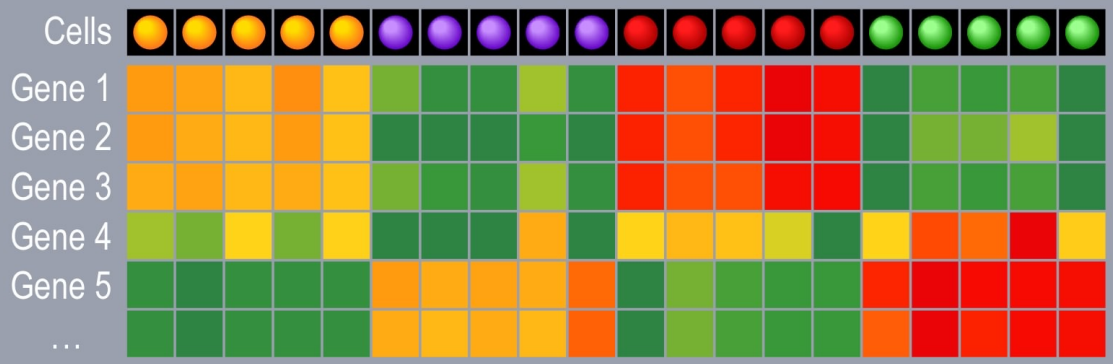
Sequencing read	Number of cycles
Read 1	26 cycles
i7 index	8 cycles
i5 index	0 cycles
Read 2	57 cycles



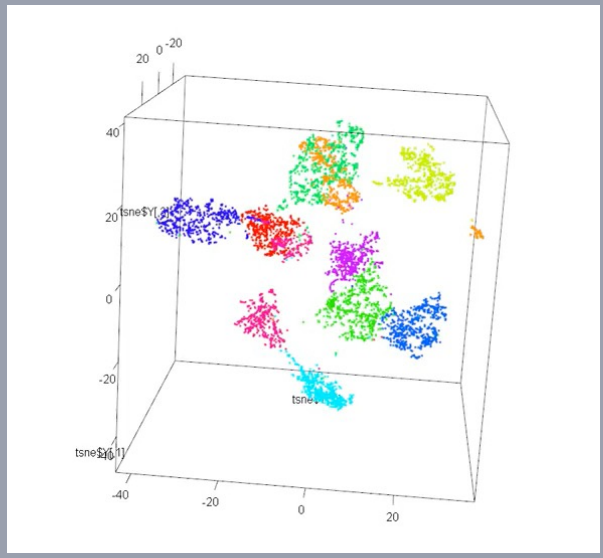
Production of a 3p gene counting expression matrix

Single Cell statistical analysis and classification

- Normalization to the Median UMI counts per cell (scaling factor)
- Matrix is log-transformed, centered and scaled per-gene (mean=0, SD=1)
- PCA analysis based on the most variables 1.000 genes
- t-SNE analysis based on 1st 10 components of the PCA-projected matrix
- k-means clustering (K=2..10) on 1st 10 components of the PCA-projected matrix
- Maximum of silhouette score as default K
- Differential expression analysis between clusters (genes markers)



3p gene counting expression matrix

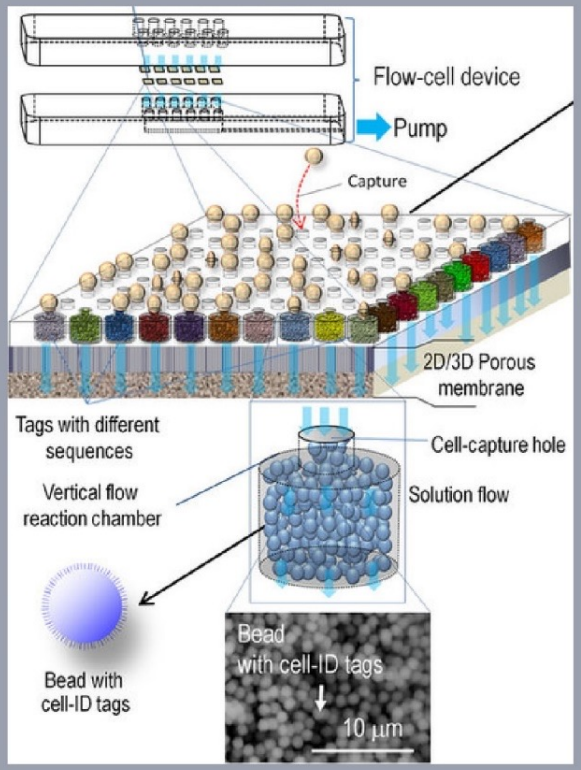


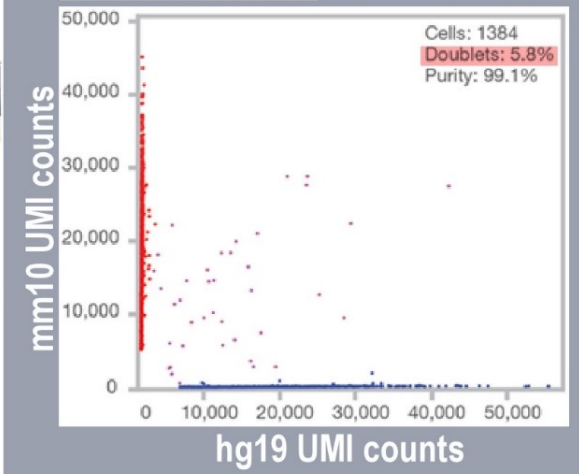
6.7 - 8.1% capture efficiency (Zheng et al., Nat.Comms., 2017)

Alternative options for high throughput single cell RNA-seq

Companies systems

- **Dolomite Bio:** Drop-seq setup
- **Hitachi:** Vertical Flow Array Chips (VFACs)
- **1CellBIO:** Kirschner's lab (*Klein et al.*) In-drop startup (Isothermal amplification)
- **Bio-Rad:** Illumina Bio-Rad SureCell™ WTA 3' Library

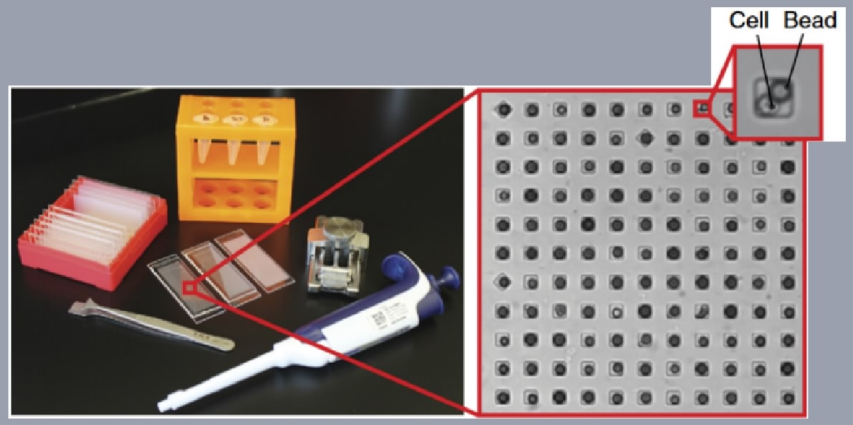




Alternative options for high throughput single cell RNA-seq

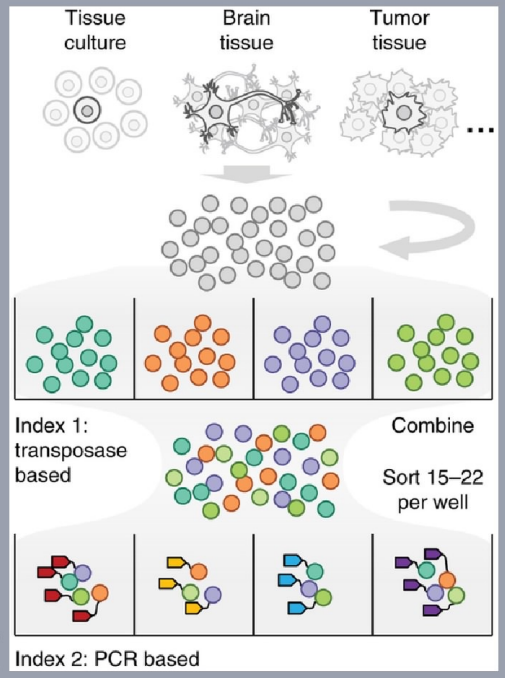
Wetlab protocol

- **Seq-Well:** portable and low cost scRNAseq in subnanoliter wells
- **SCi-seq:** combinatorial indexing (rounds of barcoding of pools of cells)



Gierahn et al., Nature Methods, feb.2017

37k in HEK, close to dropseq capture

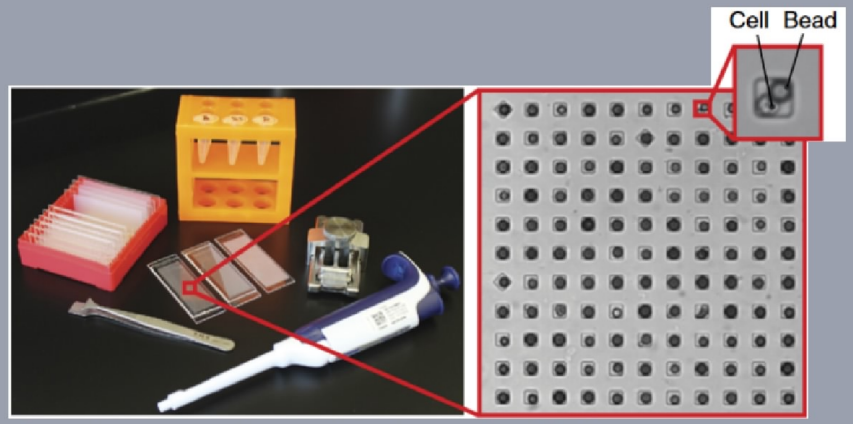


Vitak et al., Nature Methods, jan.2017

Alternative options for high throughput single cell RNA-seq

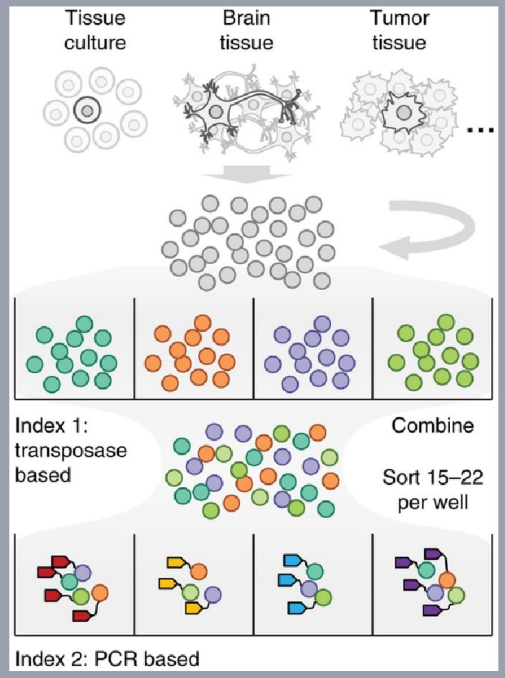
Wetlab protocol

- **Seq-Well:** portable and low cost scRNAseq in subnanoliter wells
- **SCi-seq:** combinatorial indexing (rounds of barcoding of pools of cells)



Gierahn et al., Nature Methods, feb.2017

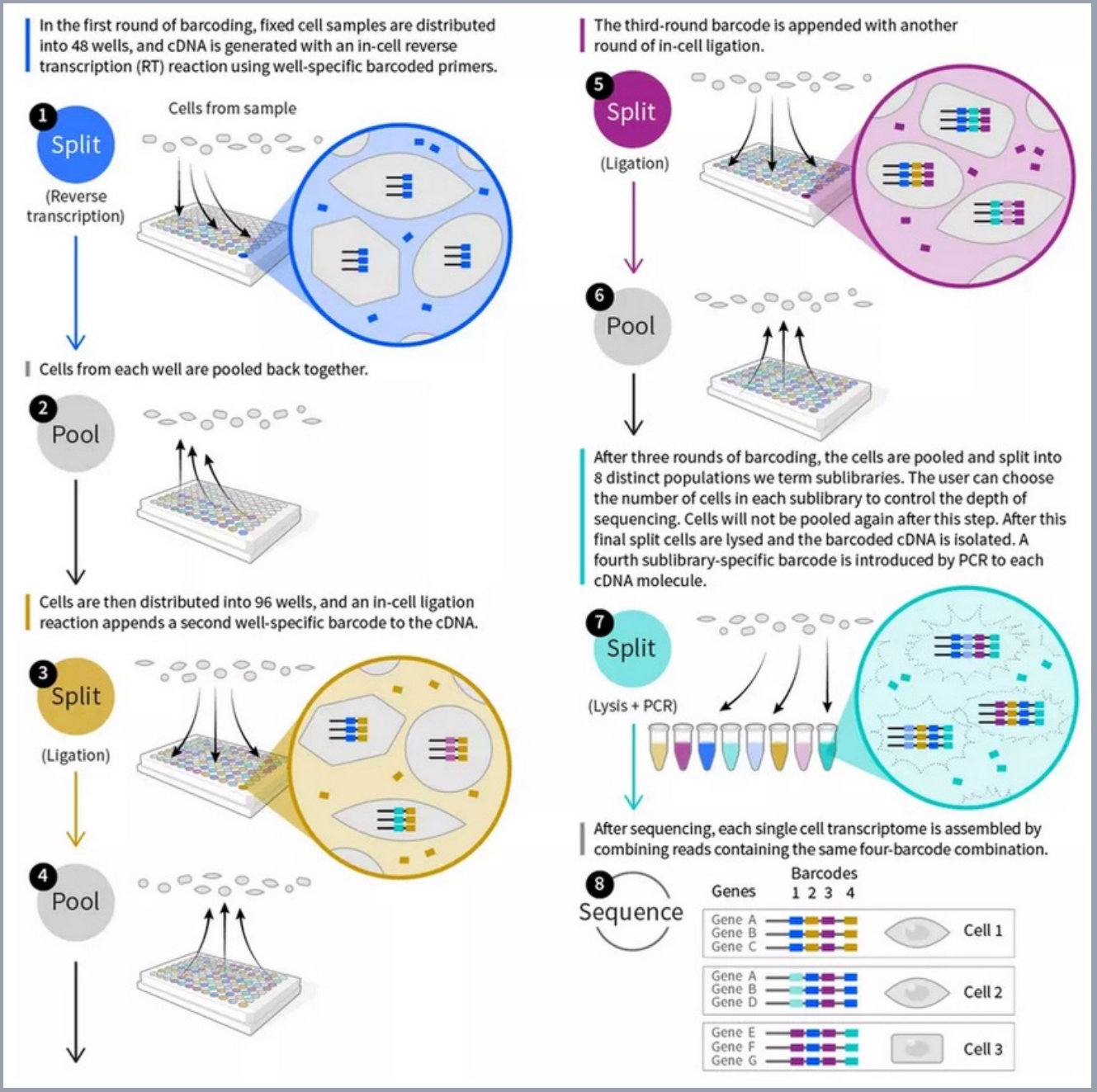
37k in HEK, close to dropseq capture



Vitak et al., Nature Methods, jan.2017

Parse BIOSCIENCES

Parsed Biosciences combinatorial indexing approach



https://www.parsebiosciences.com

Parse BIOSCIENCES

WT MINI	WT	WT MEGA
10K Cells, 12 Samples	100K Cells, 48 Samples	1M Cells, 96 Samples
Generate proof-of-concept results for a larger study, grant, or just to try the technology.	Comprehensive gene profiling across samples, replicates, or timepoints for most studies.	Expand your science or capabilities by profiling up to 1 million cells per experiment.

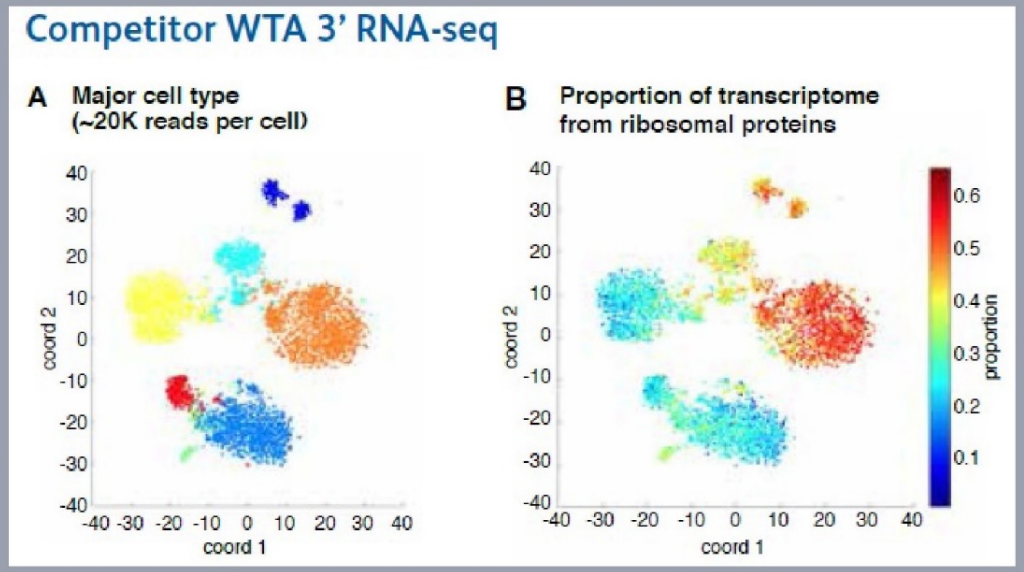
1M -> 15k€
150€ / sample

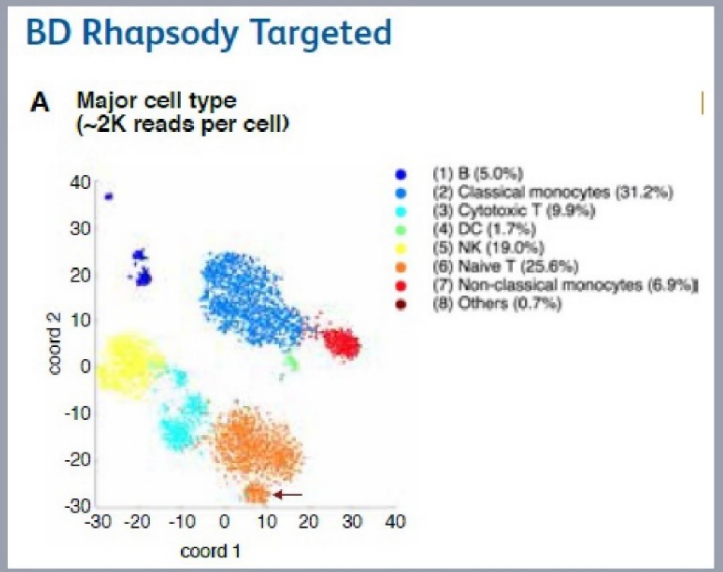
BD Rhapsody Single Cell Analysis System

Analyze 100's of genes across tens of 1000's of single cells

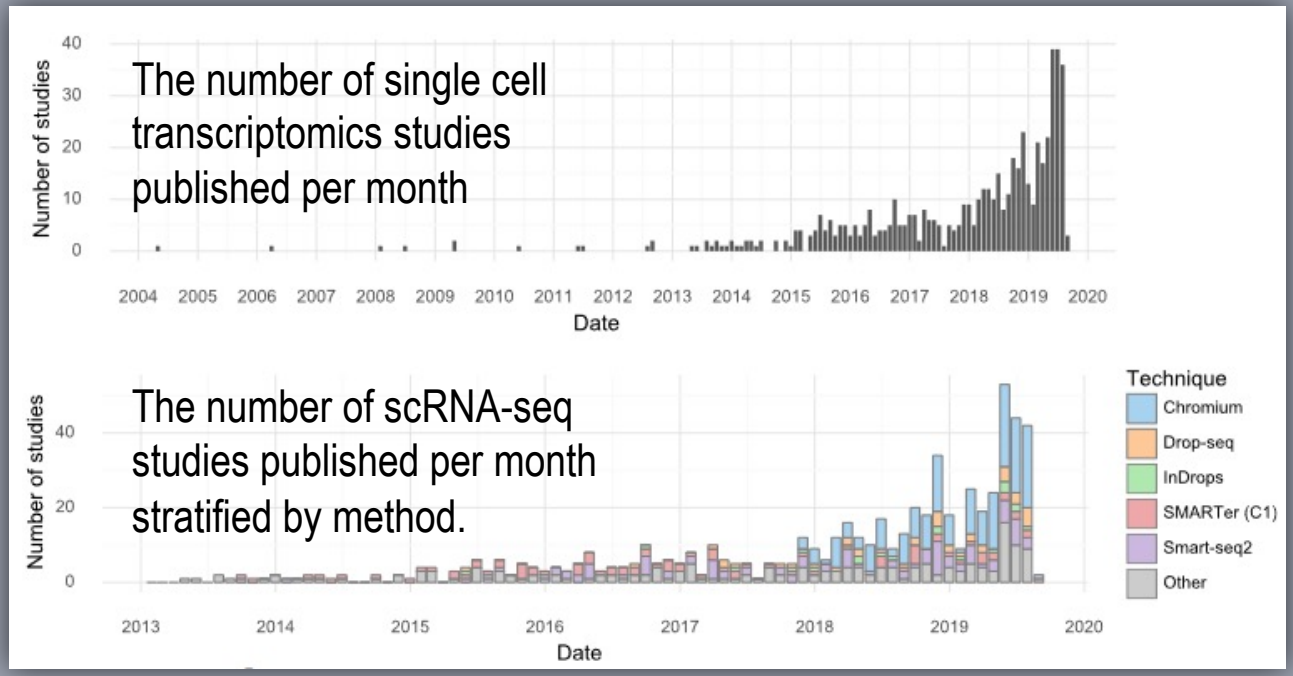
- microwells platform with barcoded beads and UMIs,
- 15.000 cells per sample,
- 99.4% count purity, minimal crosstalks between microwells,
- doublets rate close to 0% for 1.000 cells, under 5% for 15.000 cell,

- targeted assays with standard or custom gene panels:
  decrease cost of sequencing (2k reads/ cell <=> 20k reads/ cells for 10x ?)





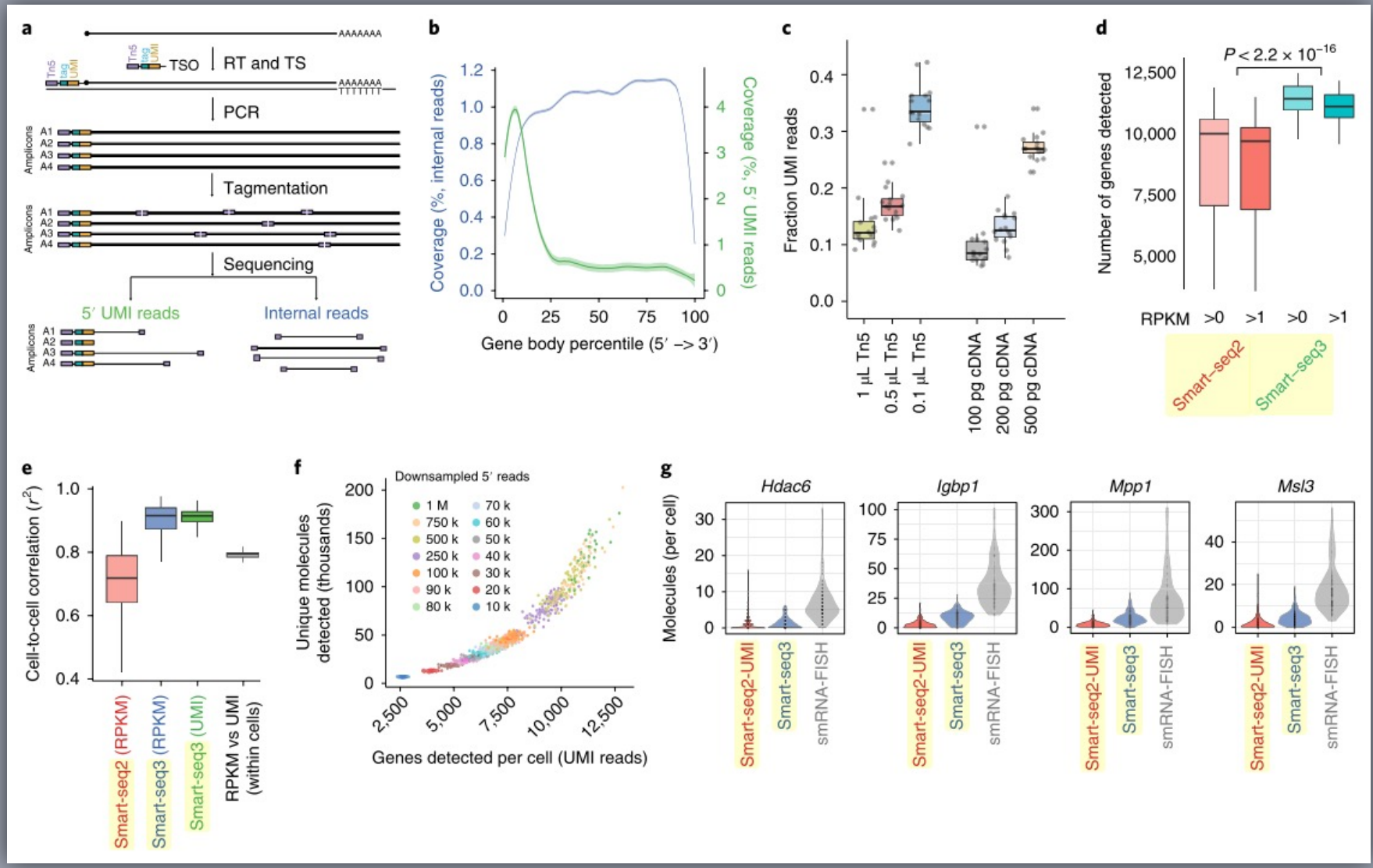
Single cell approaches in publications



Technique	Count
Chromium	655
Smart-seq2	168
SMARTer (C1)	121
Drop-seq	73
SMARTer	26
InDrops	23
CITE-seq	18
CEL-seq2	16
STRT-seq	15
Tang	15
MARS-seq	14
CEL-seq	13
Seq-Well	13
SORT-seq	12
STRT-seq (C1)	12
BD Rhapsody	9
BioMark	8
GemCode	7
ICELL8	6
Patch-seq	6
Perturb-seq	6
sc-RT-mPCR	6
MERFISH	5

A curated database reveals trends in single cell transcriptomics
Svensson et al., 2019
https://docs.google.com/spreadsheets/d/1En7-UV0k0laDiIfjFkdn7dggyR7jIk3WH8QgXaMOZF0/edit#gid=2129262271

Smart-seq3

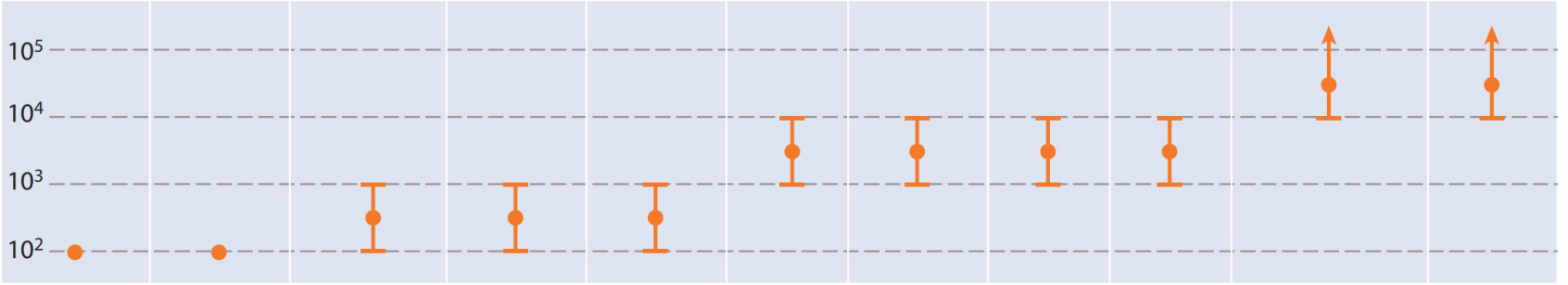


150,000 unique molecules detected per HEK293FT

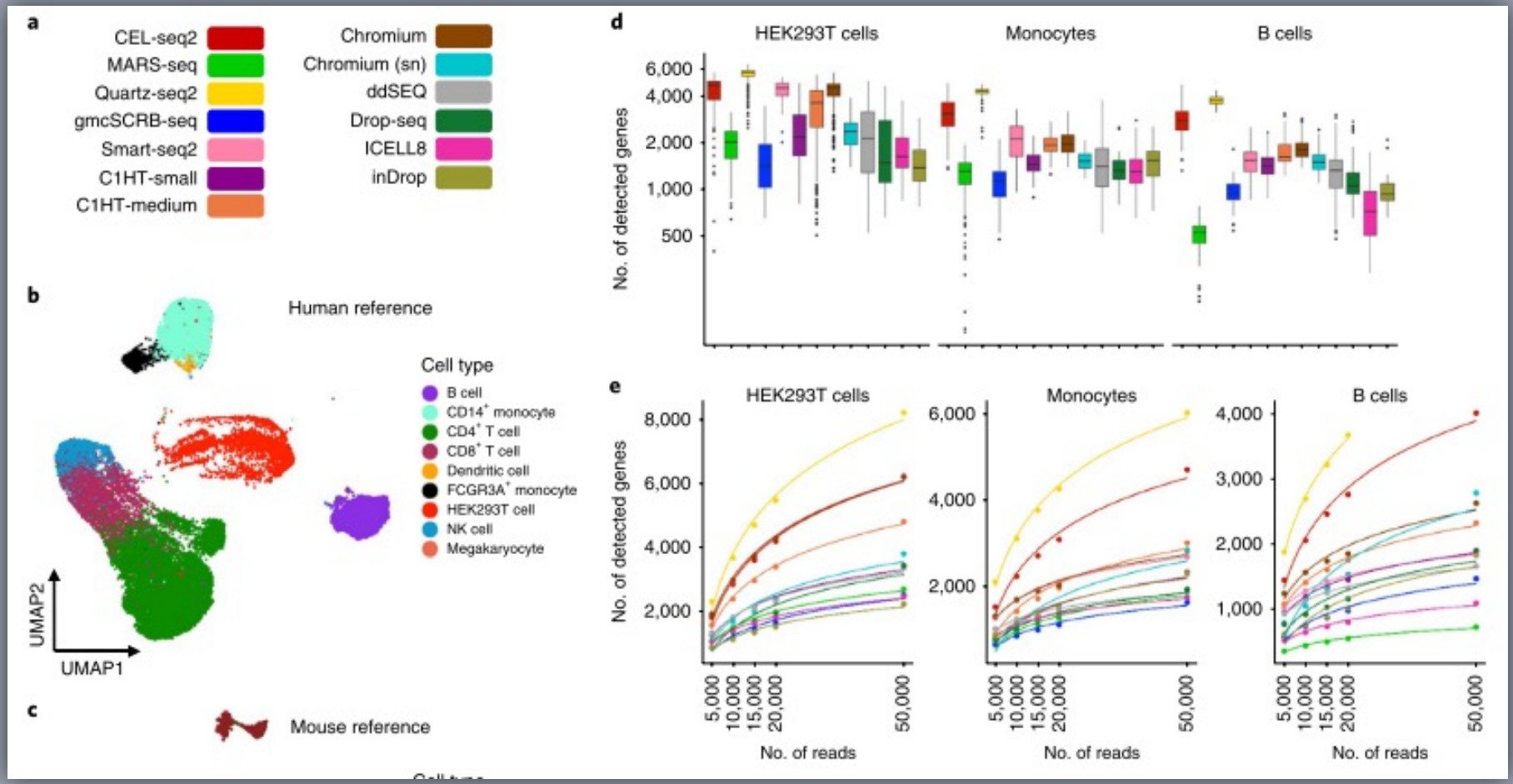
Single-cell RNA counting at allele- and isoform-resolution using Smart-seq3

Hagemann-Jensen et al., Nat.Biotech, May 2020

Experimental approaches comparison

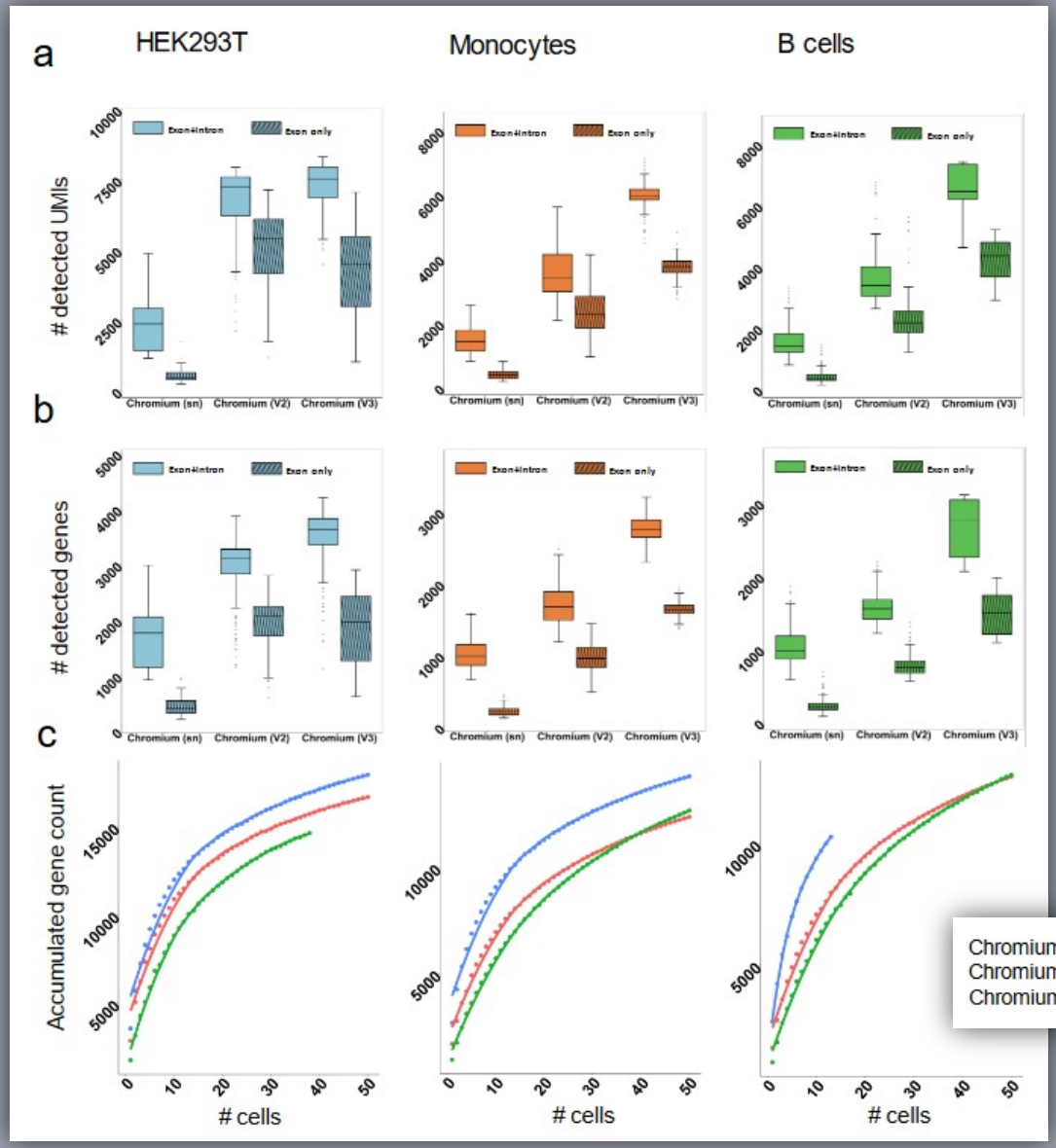
	SMART-seq2	CEL-seq2	STRT-seq	Quartz-seq2	MARS-seq	Drop-seq	inDrop	Chromium	Seq-Well	sci-RNA-seq	SPLiT-seq
Single-cell isolation	FACS, microfluidics	FACS, microfluidics	FACS, microfluidics, nanowells	FACS	FACS	Droplet	Droplet	Droplet	Nanowells	Not needed	Not needed
Second strand synthesis	TSO	RNase H and DNA pol I	TSO	PolyA tailing and primer ligation	RNase H and DNA pol I	TSO	RNase H and DNA pol I	TSO	TSO	RNase H and DNA pol I	TSO
Full-length cDNA synthesis?	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes
Barcode addition	Library PCR with barcoded primers	Barcoded RT primers	Barcoded TSOs	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers	Barcoded RT primers and library PCR with barcoded primers	Ligation of barcoded RT primers
Pooling before library?	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Library amplification	PCR	In vitro transcription	PCR	PCR	In vitro transcription	PCR	In vitro transcription	PCR	PCR	PCR	PCR
Gene coverage	Full-length	3′	5′	3′	3′	3′	3′	3′	3′	3′	3′
Number of cells per assay											

Experimental approaches comparison



Benchmarking single-cell RNA-sequencing protocols for cell atlas projects
Mereu et al., Nat.Biotech, 2020

Experimental approaches comparison



Benchmarking single-cell RNA-sequencing protocols for cell atlas projects
Mereu et al., Nat.Biotech, 2020

Experimental approaches comparison

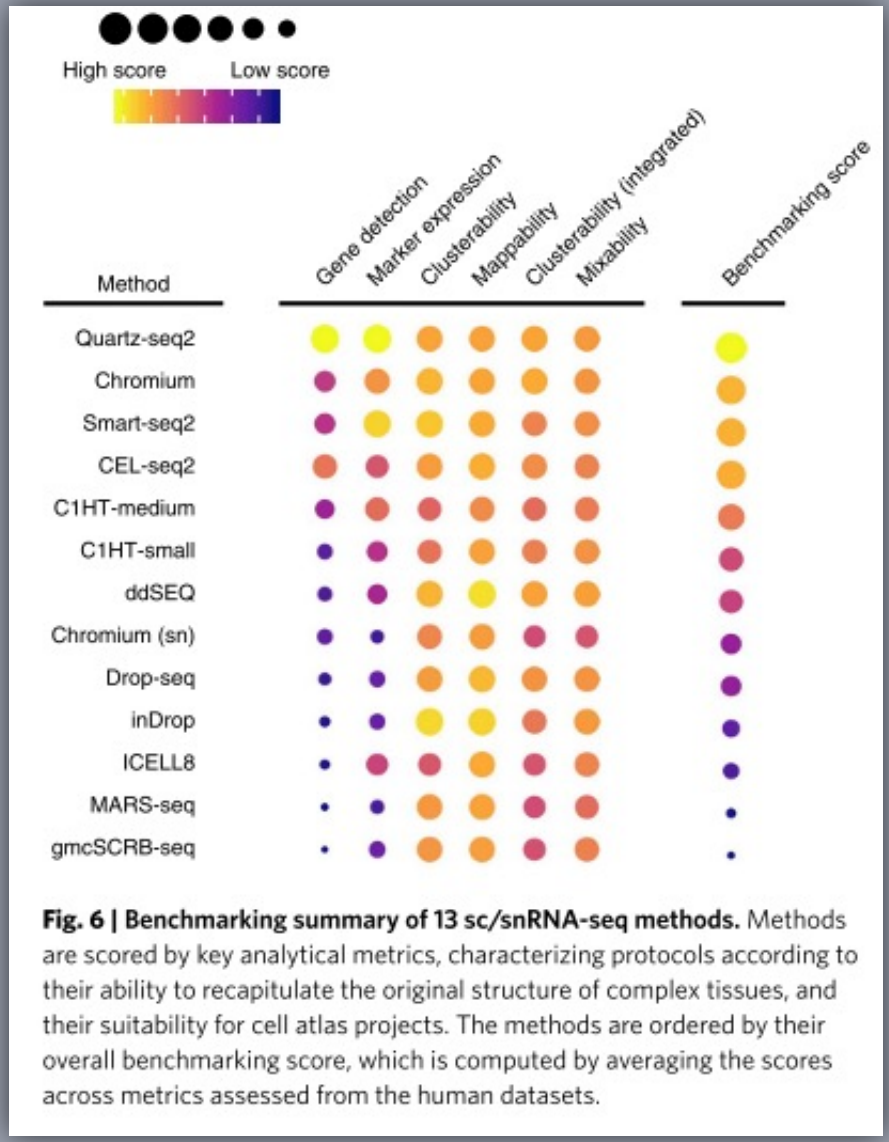
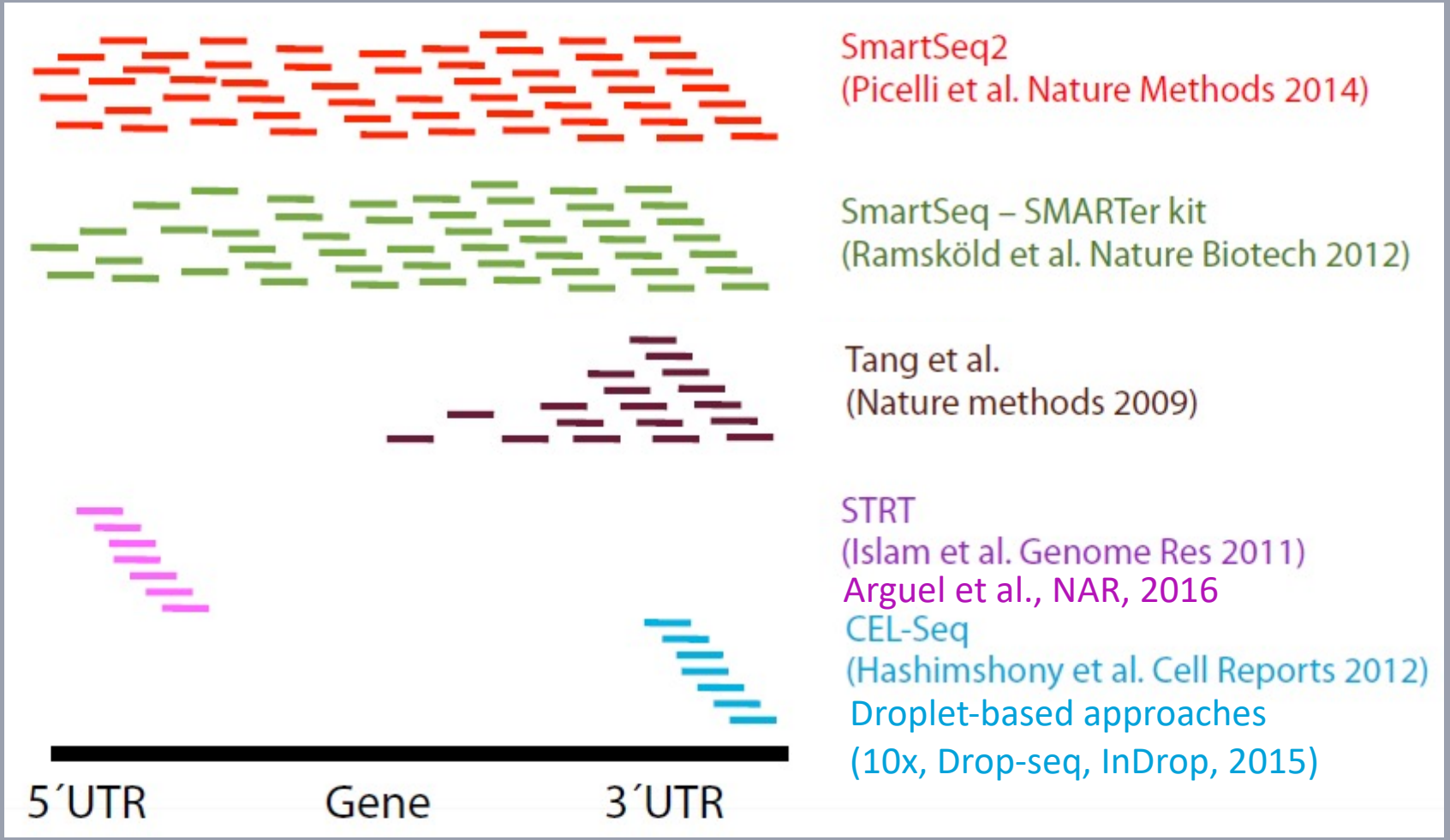


Fig. 6 | Benchmarking summary of 13 sc/snRNA-seq methods. Methods are scored by key analytical metrics, characterizing protocols according to their ability to recapitulate the original structure of complex tissues, and their suitability for cell atlas projects. The methods are ordered by their overall benchmarking score, which is computed by averaging the scores across metrics assessed from the human datasets.

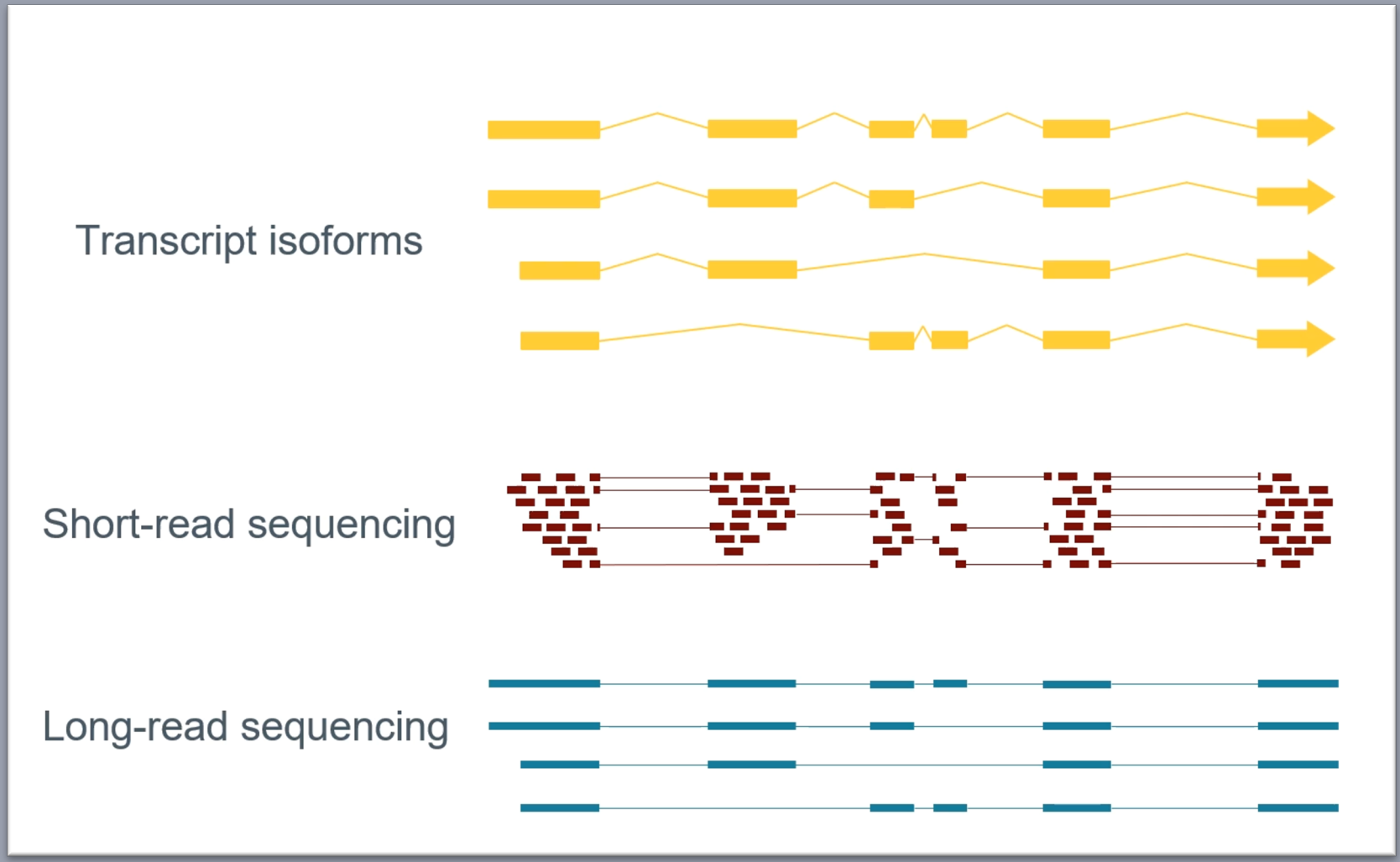
- More cells for statistics or more genes for biology ??
- Sequencing cost ??
- 5p, 3p of full length ??

Benchmarking single-cell RNA-sequencing protocols for cell atlas projects
Mereu et al., Nat.Biotech, 2020

Single Cell RNA-seq raw signal

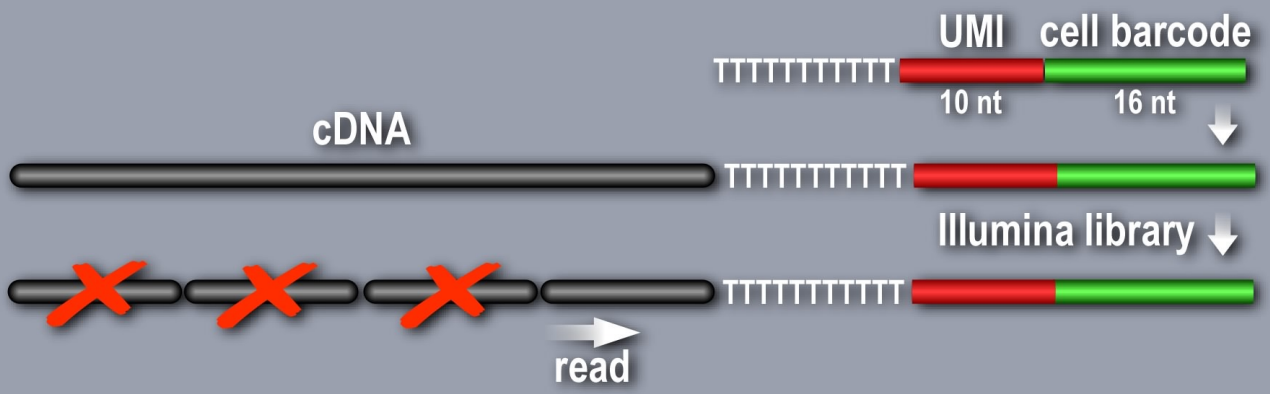


Long read sequencing identifies isoforms efficiently

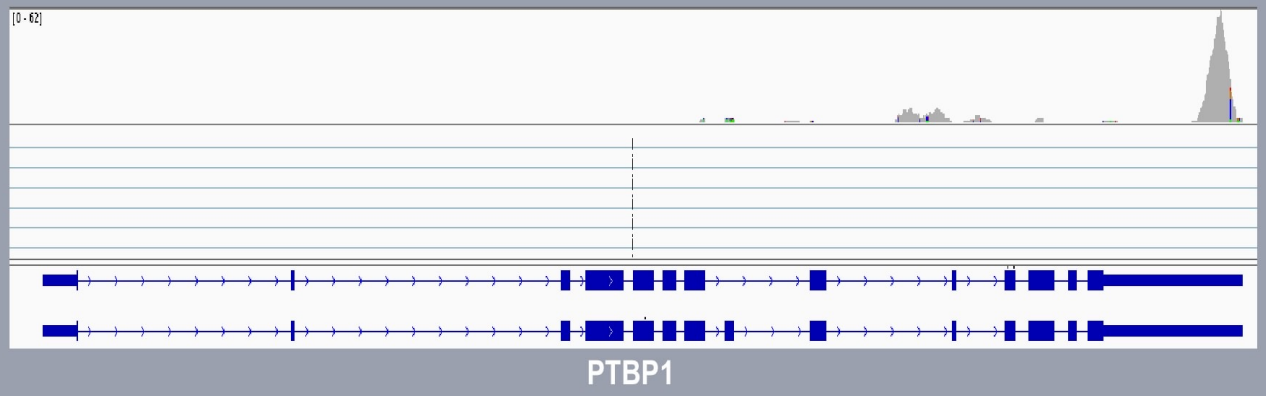


Tombacz et al., Front. Genet., 2018

Short read single cell RNA-seq



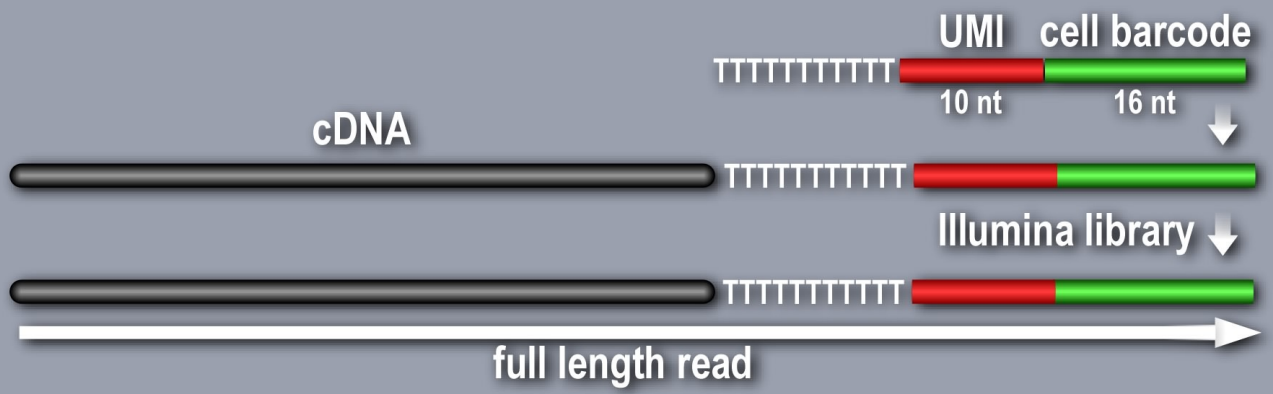
Reverse transcription, library preparation



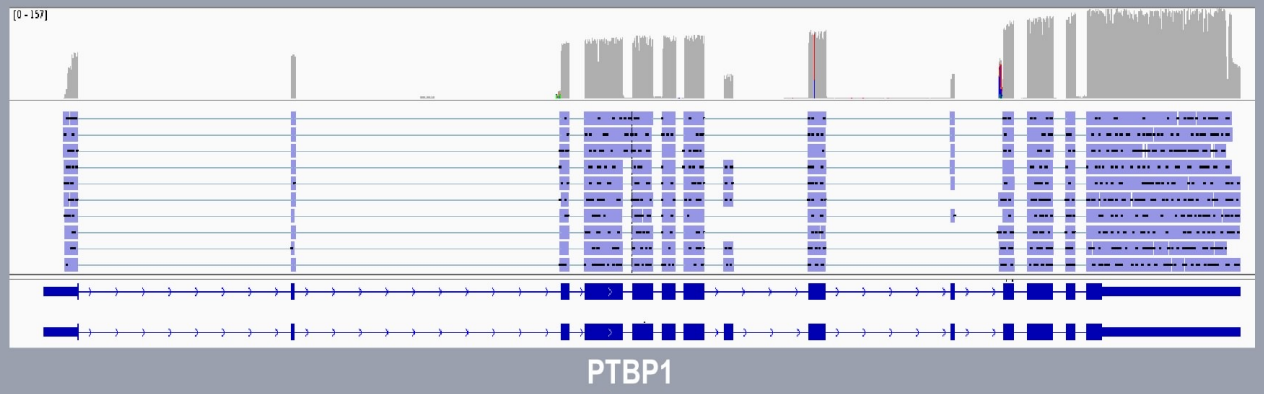
Illumina sequencing yields just a short read close to 3'

Informations on splicing, fusions, SNPs, editing, imprinting are lost

Long read single cell RNA-seq

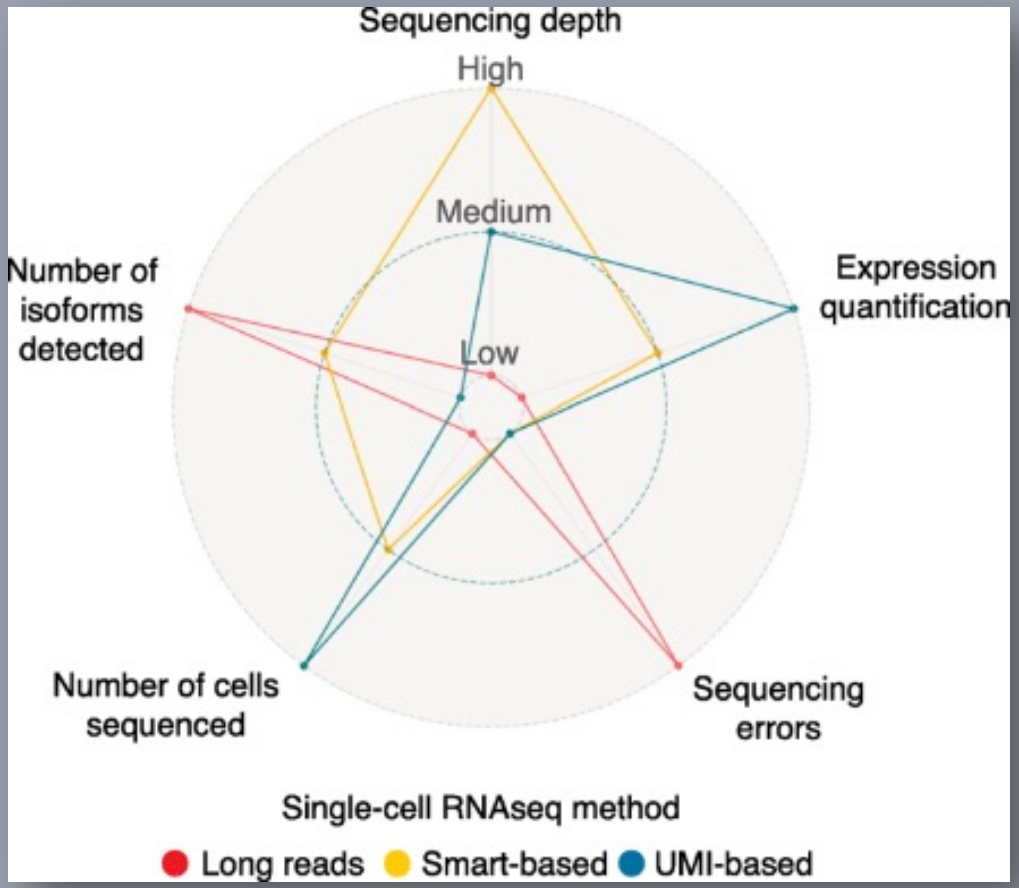


Reverse transcription, library preparation



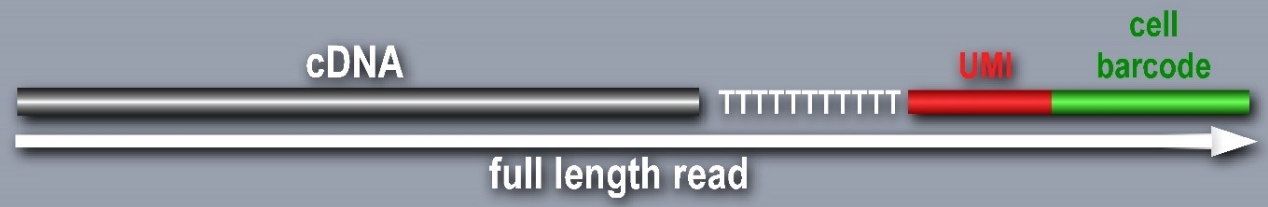
Informations on splicing, fusions, SNPs, editing, imprinting are preserved

Full-length single cell RNA-seq



Arzalluz-Luque et al. (2018)

Options for full length single cell transcriptome sequencing



PacBio Sequel II

Oxford Nanopore Technology

+++ higher accuracy	+++ high throughput (80M reads / flowcell)
--- low throughput (4M reads / SMRT)	--- lower accuracy
Price (1.400€ / SMRT)	Price (1.400€ / Flowcell)

2 challenges to tackle

(1) get enough reads to profile molecules (50k reads / cell)

(2) high accuracy for cell barcode and UMI identification

Nanopore sequencers

MinION

GridION
Runs 5 MinIon flow cells in parrallel

PromethION
1 flow cell = 6 MinIon
Cost / read 3x lower

Library preparation is straightforward

cDNA or genomic DNA ➡ ligate Nanopore adapters ➡ sequence

February 2017, Linnarsson's lab

Karlsson and Linnarsson *BMC Genomics* (2017) 18:126
DOI 10.1186/s12864-017-3528-6

BMC Genomics

RESEARCH ARTICLE Open Access

Single-cell mRNA isoform diversity in the mouse brain

Kasper Karlsson[1] and Sten Linnarsson[2*]

Abstract

Background: Alternative mRNA isoform usage is an important source of protein diversity in mammalian cells. This phenomenon has been extensively studied in bulk tissues, however, it remains unclear how this diversity is reflected in single cells.

Results: Here we use long-read sequencing technology combined with unique molecular identifiers (UMIs) to reveal patterns of alternative full-length isoform expression in single cells from the mouse brain. We found a surprising amount of isoform diversity, even after applying a conservative definition of what constitutes an isoform. Genes tend to have one or a few isoforms highly expressed and a larger number of isoforms expressed at a low level. However, for many genes, nearly every sequenced mRNA molecule was unique, and many events affected coding regions suggesting previously unknown protein diversity in single cells. Exon junctions in coding regions were less prone to splicing errors than those in non-coding regions, indicating purifying selection on splice donor and acceptor efficiency.

Conclusions: Our findings indicate that mRNA isoform diversity is an important source of biological variability also in single cells.

Keywords: Alternative isoform usage, Single-cell RNA sequencing, STRT, PacBio, Long read sequencing, UMI, Oligodendrocytes

Feb.2017:

We selected six single cells for which cDNA was available from an earlier experiment … was used for PacBio sequencing. The cDNA had been produced with the STRT method adapted to the Fluidigm C1 instrument for single cell RNA sequencing

July 2017, Vollmer's lab

Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells

Ashley Byrne, Anna E. Beaudin, Hugh E. Olsen, Miten Jain, Charles Cole, Theron Palmer, Rebecca M. DuBois, E. Camilla Forsberg, Mark Akeson & Christopher Vollmers

Nature Communications **8**,
Article number: 16027 (2017)
doi:10.1038/ncomms16027

Received: 24 April 2017
Accepted: 23 May 2017
Published: 19 July 2017

Understanding gene regulation and function requires a genome-wide method capable of capturing both gene expression levels and isoform diversity at the single-cell level. Short-read RNAseq is limited in its ability to resolve complex isoforms because it fails to sequence full-length cDNA copies of RNA molecules. Here, we investigate whether RNAseq using the long-read single-molecule Oxford Nanopore MinION sequencer is able to identify and quantify complex isoforms without sacrificing accurate gene expression quantification. After benchmarking our approach, we analyse individual murine B1a cells using a custom multiplexing strategy. We identify thousands of unannotated transcription start and end sites, as well as hundreds of alternative splicing events in these B1a cells. We also identify hundreds of genes expressed across B1a cells that display multiple complex isoforms, including several B cell-specific surface receptors. Our results show that we can identify and quantify complex isoforms at the single cell level.

July 2017:

To test this, we used our ONT RNAseq approach to analyse seven individual mouse B1a cells and compared it with the standard Illumina RNAseq approach. To this end, we FACS-sorted single B1a cells into individual wells containing lysis buffer and amplified cDNA from each individual cell using a modified Smartseq2

Juin 2018:

R2C2: Improving nanopore read accuracy enables the sequencing of highly-multiplexed full-length single-cell cDNA , Volden et al., Vollmers's lab, bioRxiv, 2018

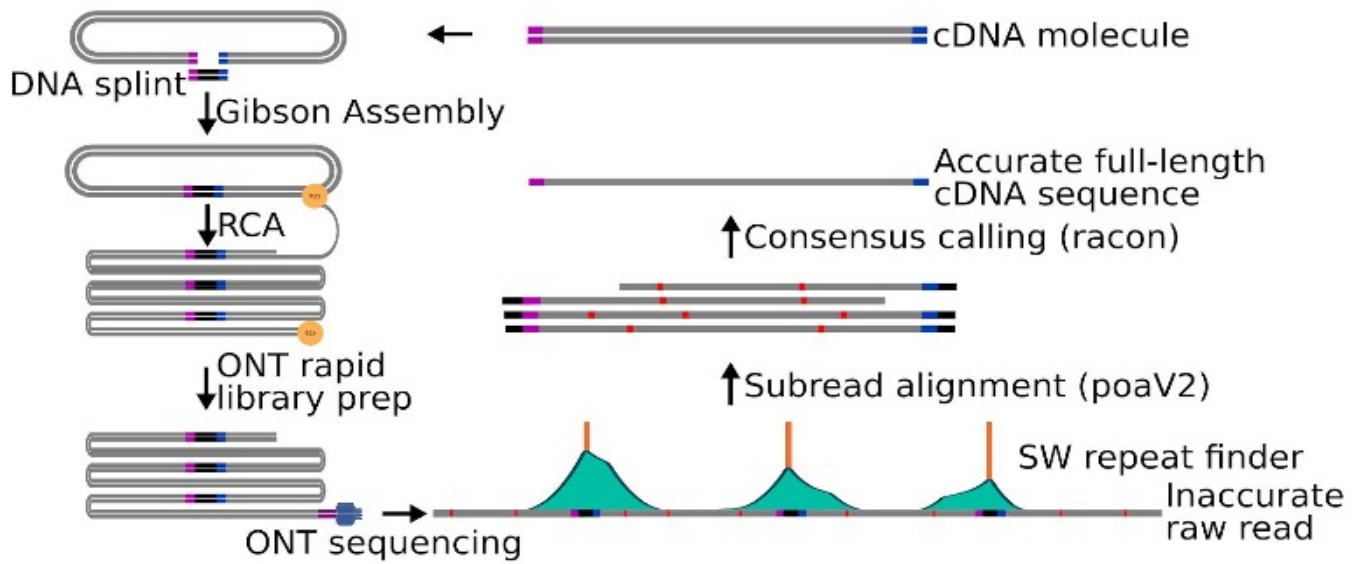


Fig. 1: R2C2 method overview. A) *cDNA is circularized using Gibson Assembly, amplified using RCA, and sequenced using the ONT MinION. The resulting raw reads are split into subreads containing full-length or partial cDNA sequences, which are combined into an accurate consensus sequences using our C3POa workflow which relies on a custom algorithm to detect DNA splints as well as poaV2 and racon.*

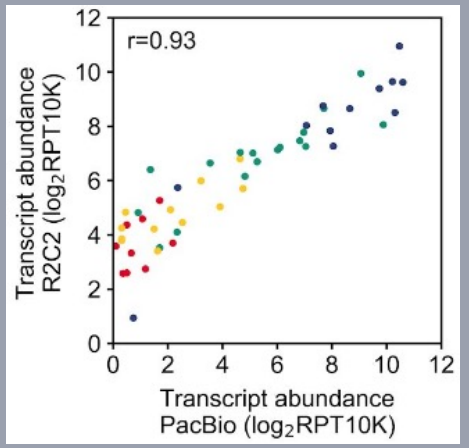
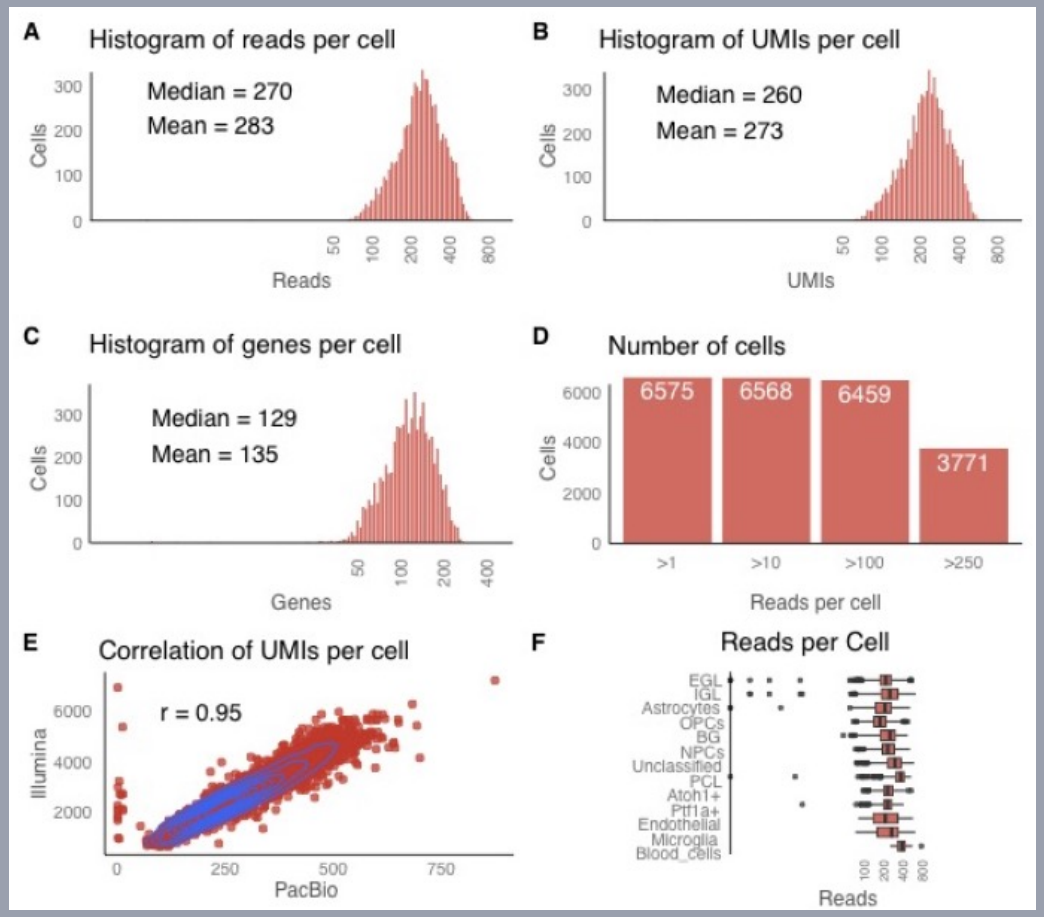


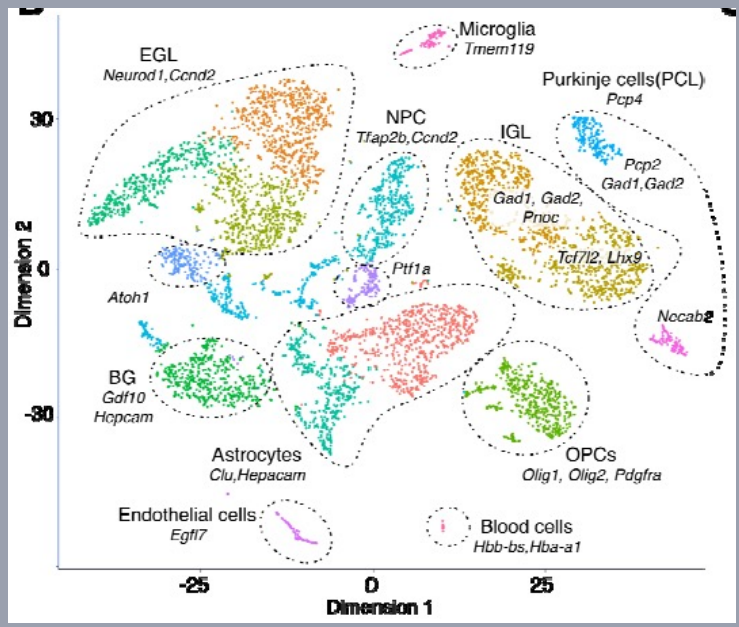
Plate of 96 B cells, Smart-seq2 modified protocol with Rolling Circle Amplification (RCA) and sequence with 1D R9.5 ONT flowcells. Good correlation with PacBio profiling, R2C2 can be easily adapted to any RNAseq library preparation protocol (10x, smart-seq2, drop-seq)

October 2018, Tilger's Lab, Nature Biotechnology, ScISOr-Seq

Single-cell isoform RNA sequencing characterizes isoforms in thousands of cerebellar cells

Ishaan Gupta, Paul G Collier, Bettina Haase, Ahmed Mahfouz, Anoushka Joglekar, Taylor Floyd, Frank Koopmans, Ben Barres, August B Smit, Steven A Sloan, Wenjie Luo, Olivier Fedrigo, M Elizabeth Ross & Hagen U Tilgner





Oct.2018:

6.627 mouse P1 cerebellum cells

10xGenomics sample

23 SMRT PacBIo flowcells

5.2M reads generated

260 UMIs per cell out of the

3.875 UMIs identified in Illumina.

scNaUMI-seq：11-2019 (Biorxiv), 08-2020 (Nature Comm.)

nature > nature communications > articles > article

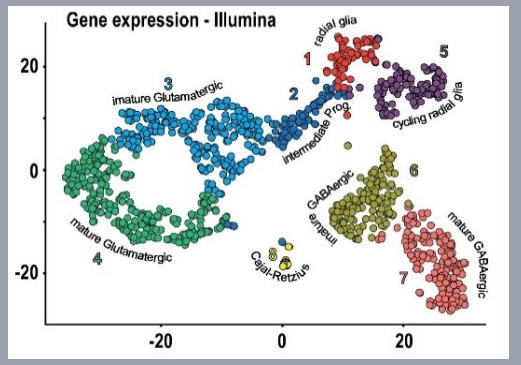
Article | Open Access | Published: 12 August 2020

High throughput error corrected Nanopore single cell transcriptome sequencing

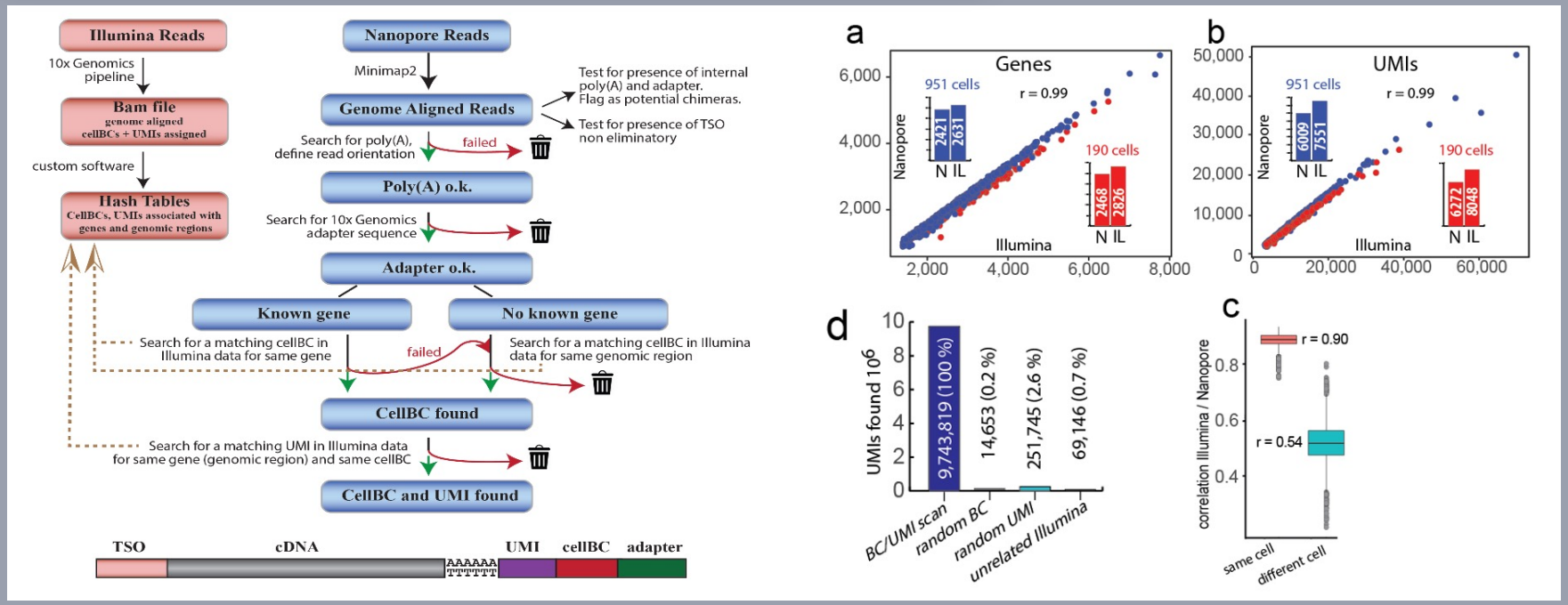
Kevin Lebrigand, Virginie Magnone, Pascal Barbry & Rainer Waldmann

Nature Communications **11**, Article number: 4025 (2020) | Cite this article

3976 Accesses | **61** Altmetric | Metrics

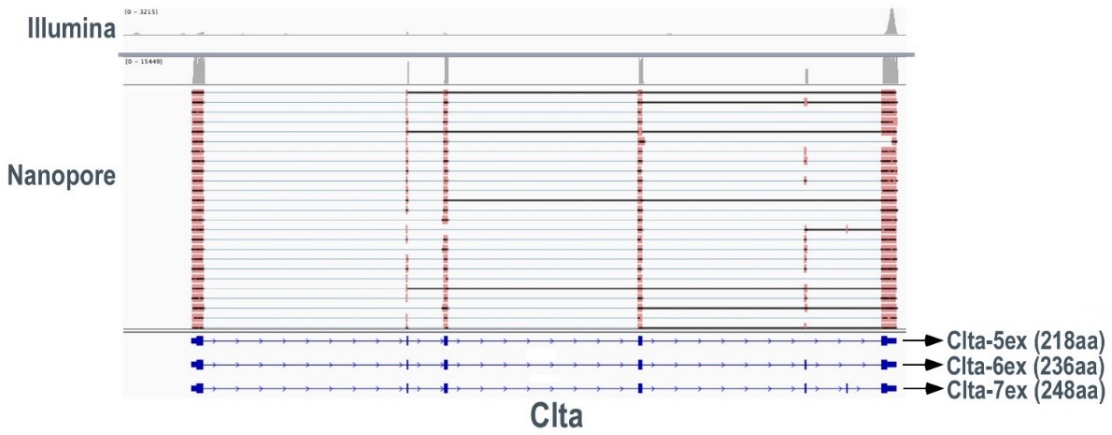


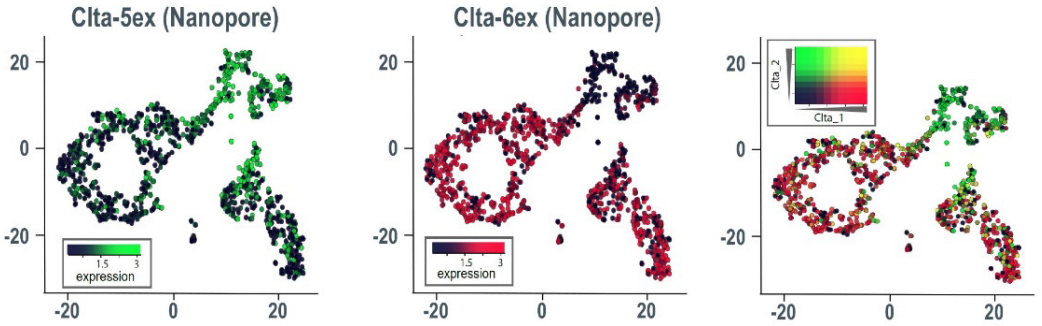
1,141 mouse brain cells

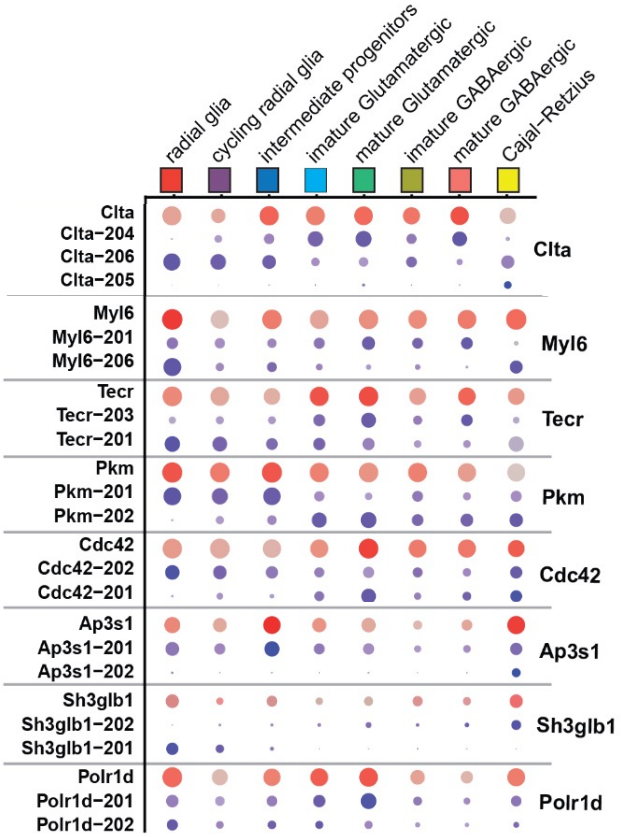


https://github.com/ucagenomix/sicelore

Long read sequencing reveals diversity







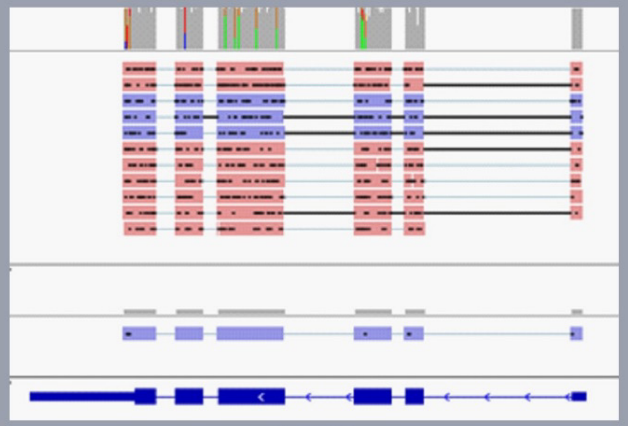
scNaUMI-seq

Long read sequencing allows SNV calling in full length molecule

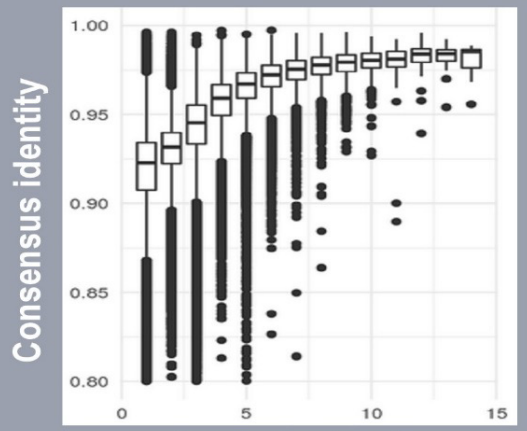
- Multiple alignment of reads to define consensus sequence (if >10 reads, best 10 reads used)
- racon polishing using all reads

Reads from the same ACTB molecule 92% identity each

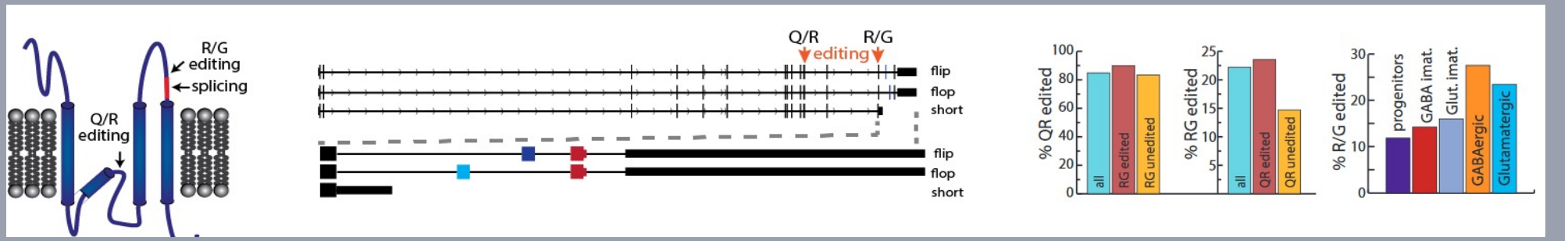
Consensus sequence 99.2% identity



ACTB



Reads / molecule



Summary

- A lot of single cell protocol available
  Advantages and drawbacks need to be known (cost ??)
- More genes (plate-based) versus more cells (droplet-based) ?
  Biological questions to address must drive protocol selection
- 3p versus full length protocols
  Smart-seq not the only option for full length heterogeneity exploration