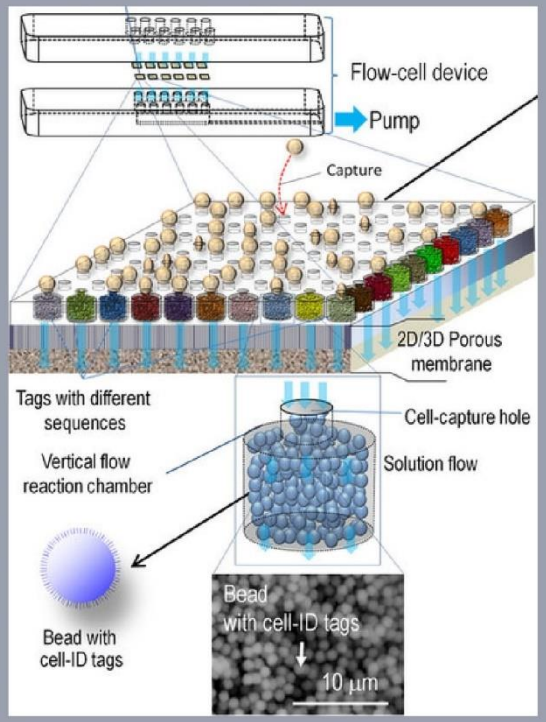
Singel Cell RNA-seq: Technologies and Experimental Approaches

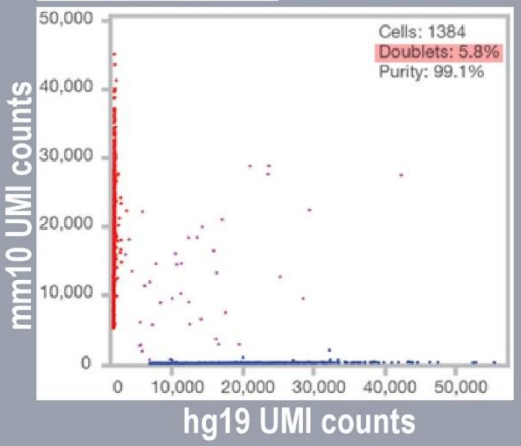
Kévin Lebrigand
UCAGenomix, Nice-Sophia-Antipolis
lebrigand@ipmc.cnrs.fr
@kevinlebrigand
Roscoff, 18 juin 2018

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

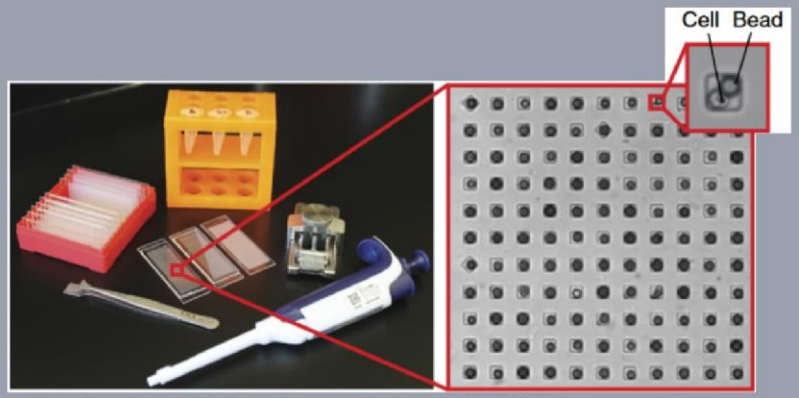




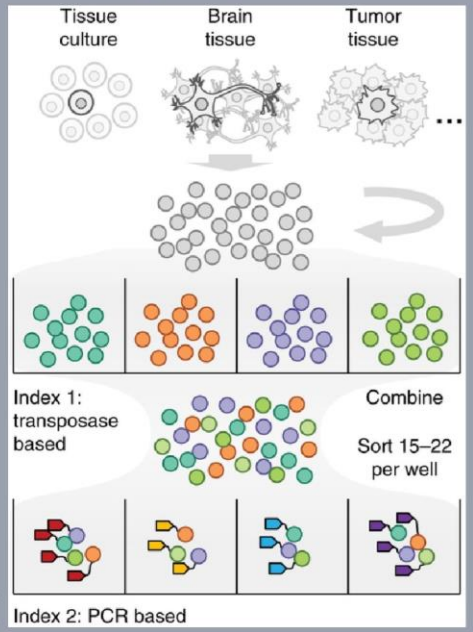
Alternative options for high throughput single cell RNA-seq

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



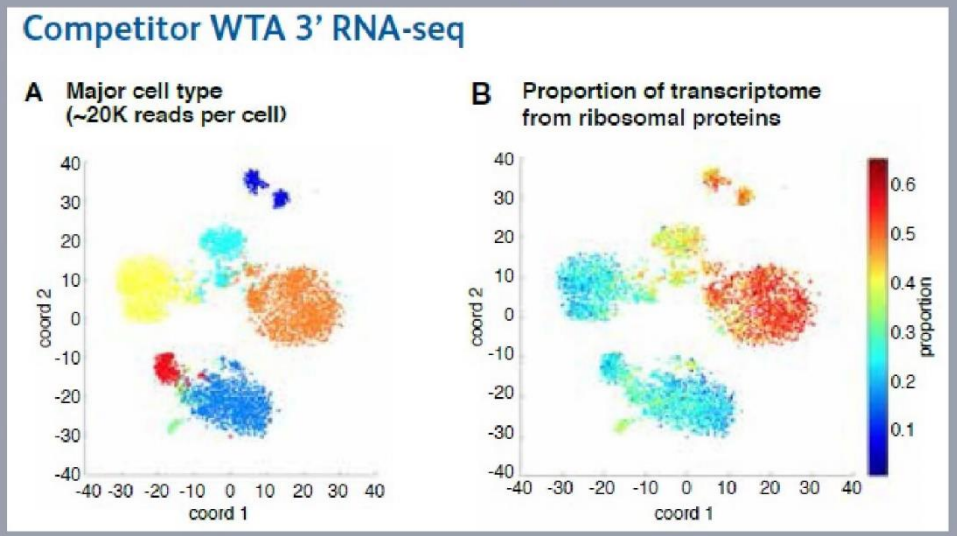
Vitak et al., Nature Methods, jan.2017

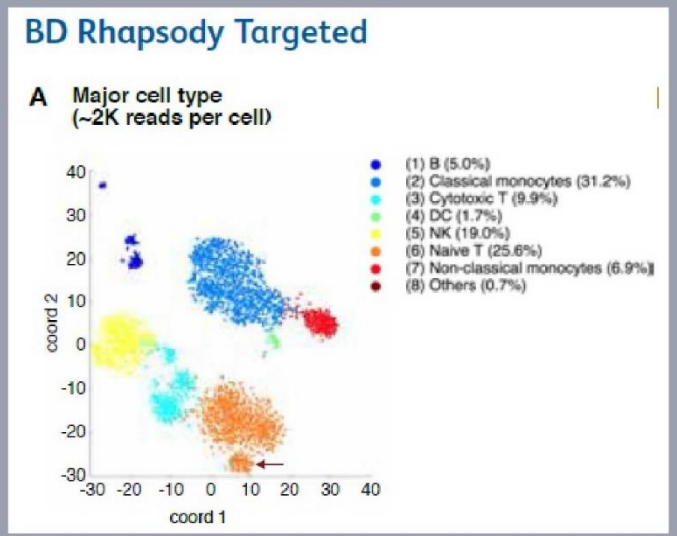
BD Rhapsody Single Cell Analysis System

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 ?)



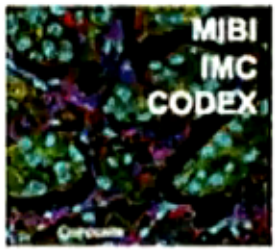


Approaches for Spatial Genomics

Direct Measurement

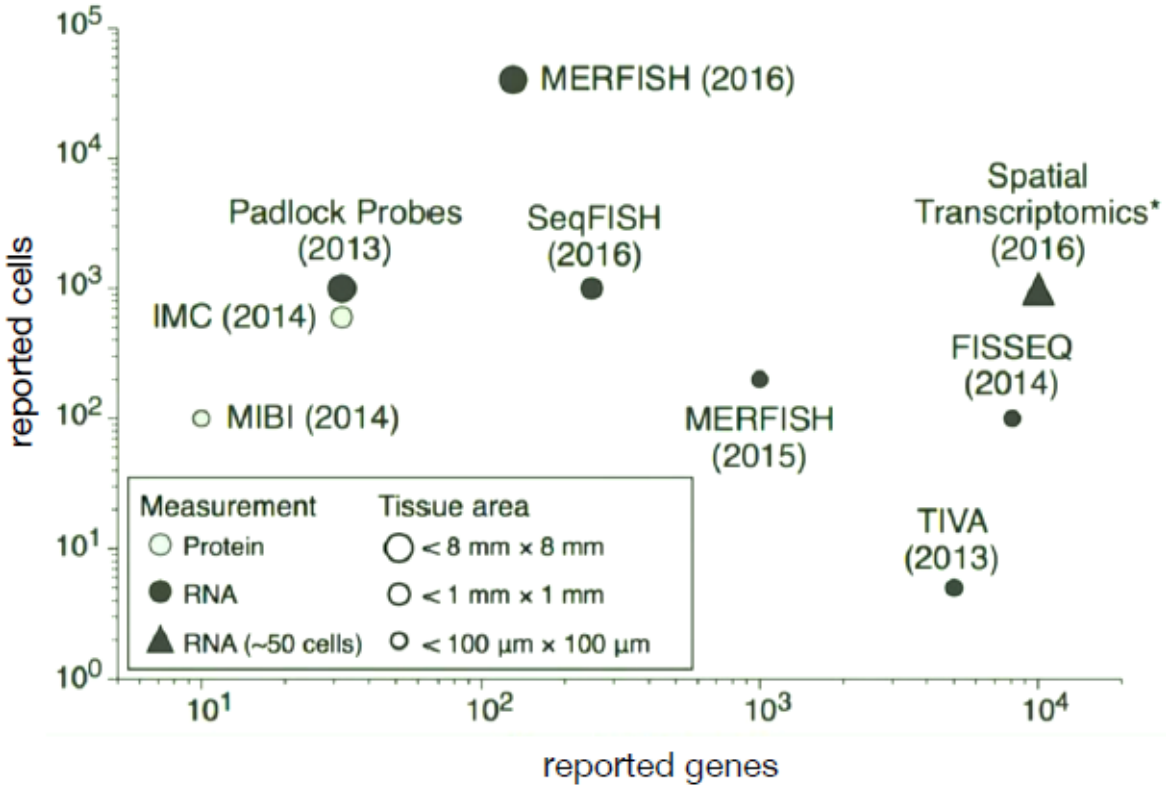
multiplex RNA fISH

in situ multi-plex protein localization



In situ sequencing

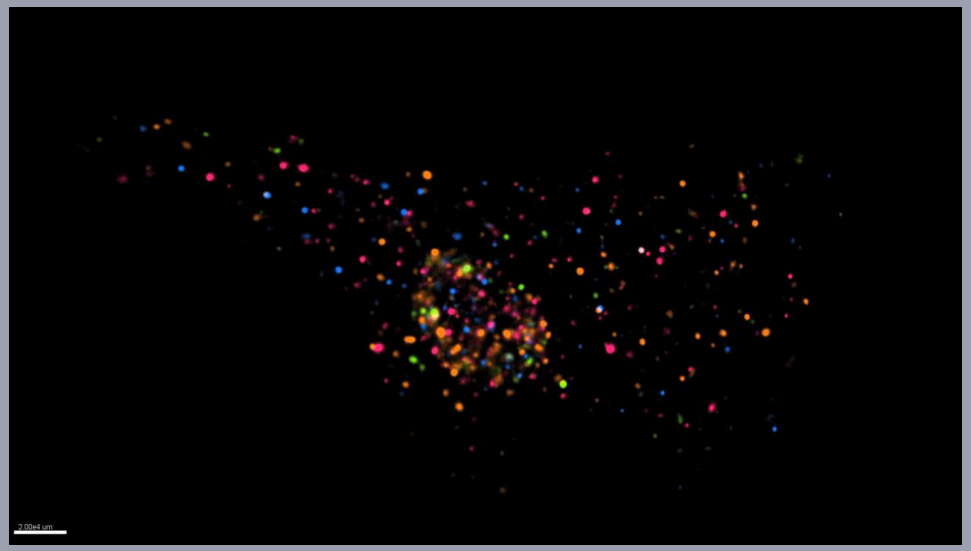
Spatial Technologies



FISSEQ, Lee et al. Science, 2014

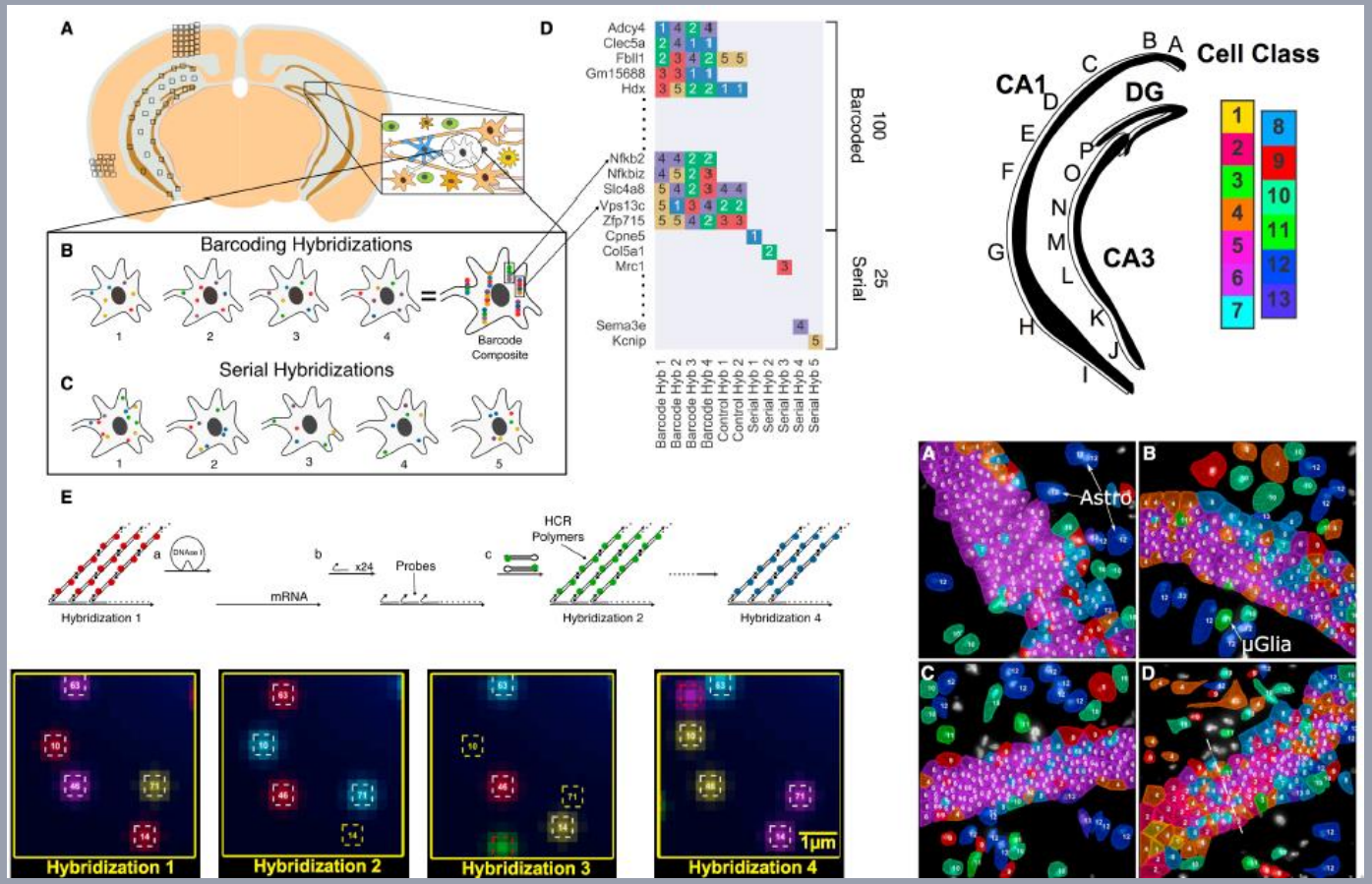
- **Sequencing in situ (FISSEQ):** yields precise information on both cellular and subcellular localisation of transcripts.
  Lee at al., Science, Mar.2014

Cross-linked 3D RNA-seq library in cells



Based on **SOLID** chemistry and **sequencing by ligation**, you can get info on cellular and sub-cellular localisation of transcripts
2 days experiment for reads of 25 bases long

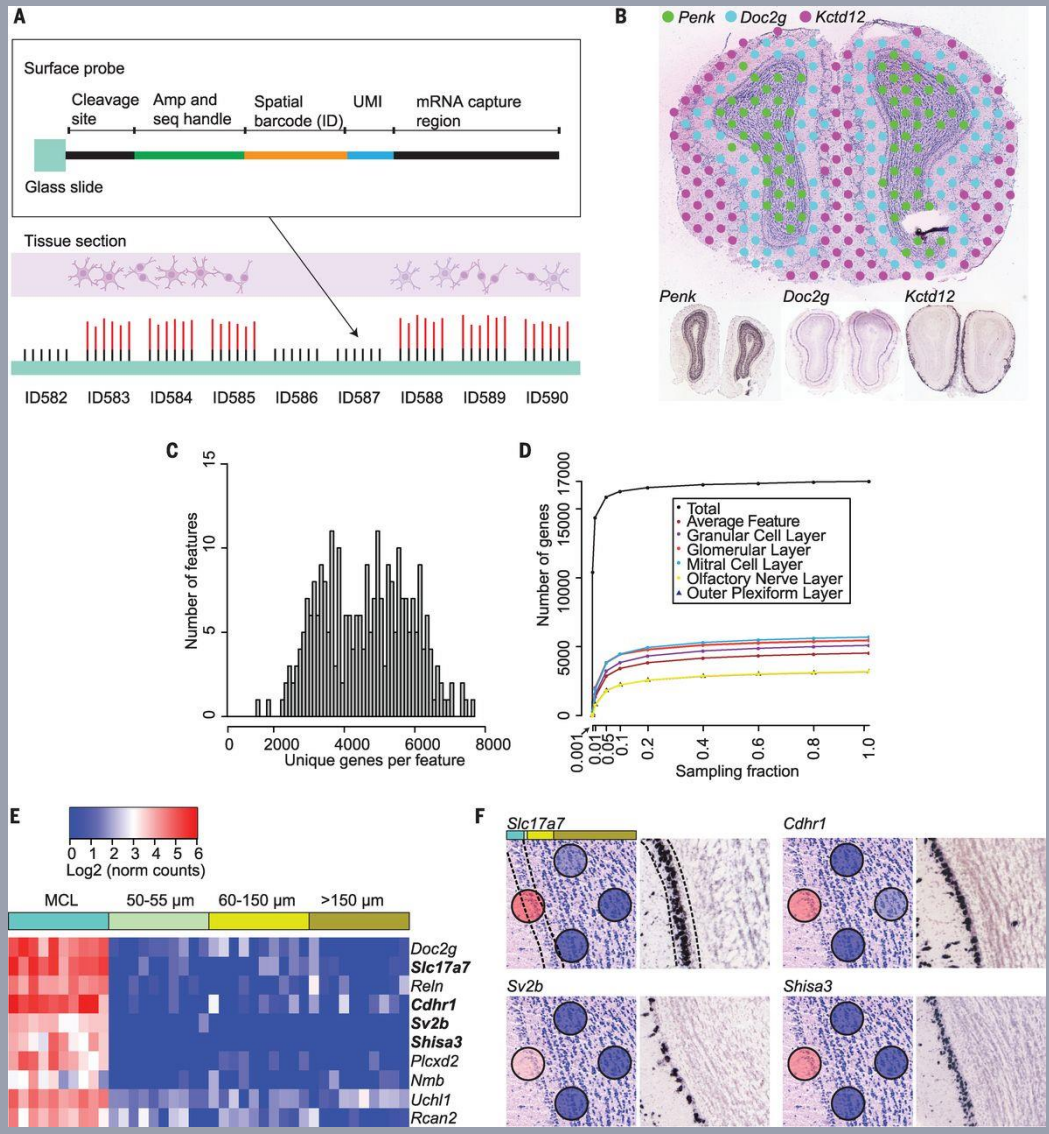
- **Highly multiplexed in-situ hybridization (Amplified seqFISH)**
  in situ 3D multiplexed imaging method



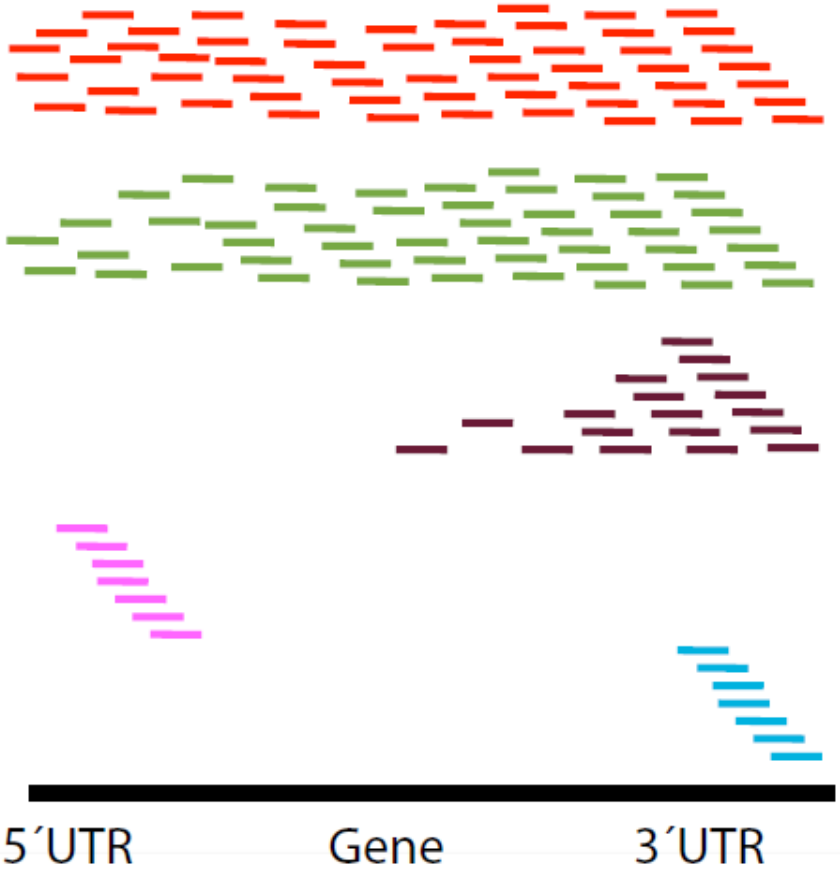
4 fluorophores and 4 rounds of hyridization can discriminate 4^4 genes = 256 genes

Spatial transcriptomics, Stahl et al. Science, 2016

- **Spatial RNA-seq with position-indexed RT primer arrays on slides (spatial transcriptomics)**
  200 million oligos, 1007 features (100µm), ~5.000 genes per feature.
  Stahl et al., Science, Aug.2016



Single Cell RNA-seq methods



SmartSeq2
(Picelli et al. Nature Methods 2014)

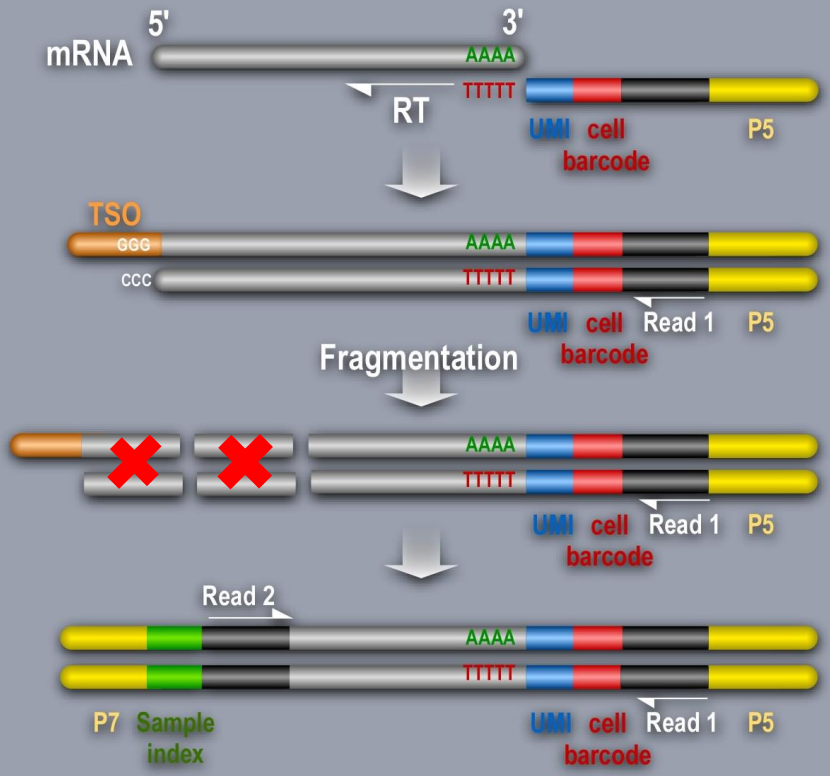
SmartSeq – SMARTer kit
(Ramsköld et al. Nature Biotech 2012)

Tang et al.
(Nature methods 2009)

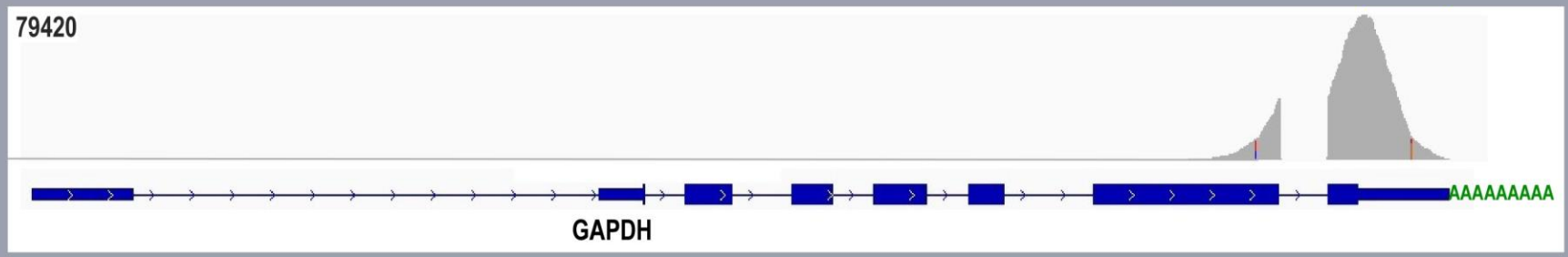
STRT
(Islam et al. Genome Res 2011)
Arguel et al., NAR, 2016
CEL-Seq
(Hashimshony et al. Cell Reports 2012)

Droplet-based approaches
(Drop-seq, InDrop, 2015)

10xGenomics 3p selective signal



Read the 5' of the most 3' fragment



Sequencing with long reads

Advantages of long reads: whole length information

- alternative splicing identification, exon retention event, isoforms counting,
- SNPs, editing events,
- potentially fusion transcripts identification,

Issues with Nanopore sequencing:

- low percentage identity (~90%) induce issues in cell barcode and UMI identification,
- low sequencing thoughtput (cost+++)
- protocol optimization, fragmentation--, Nanopore needs 100's to 1000 more input material (PCR+++ > biais+++)

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

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

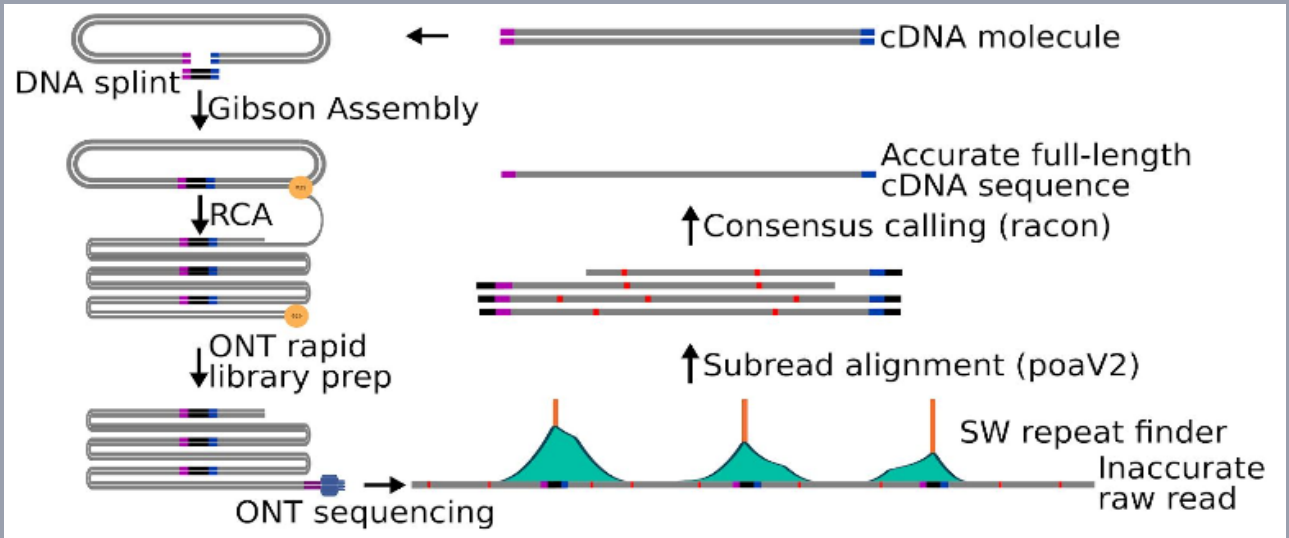


Fig. 1: R2C2 method overview. A) *cDNA is circularized using Gibson Assembly, amplified using RCA, and sequenced using the ONT MinION. The resulting raw reads are split into subreads containing full-length or partial cDNA sequences, which are combined into an accurate consensus sequences using our C3POa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

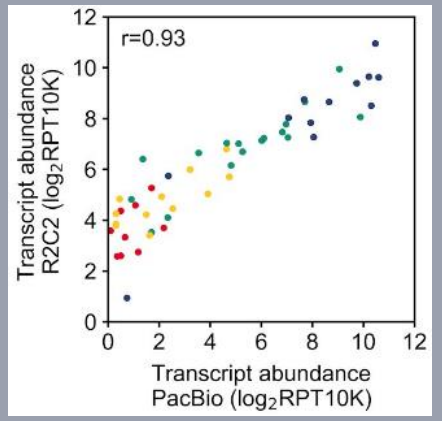
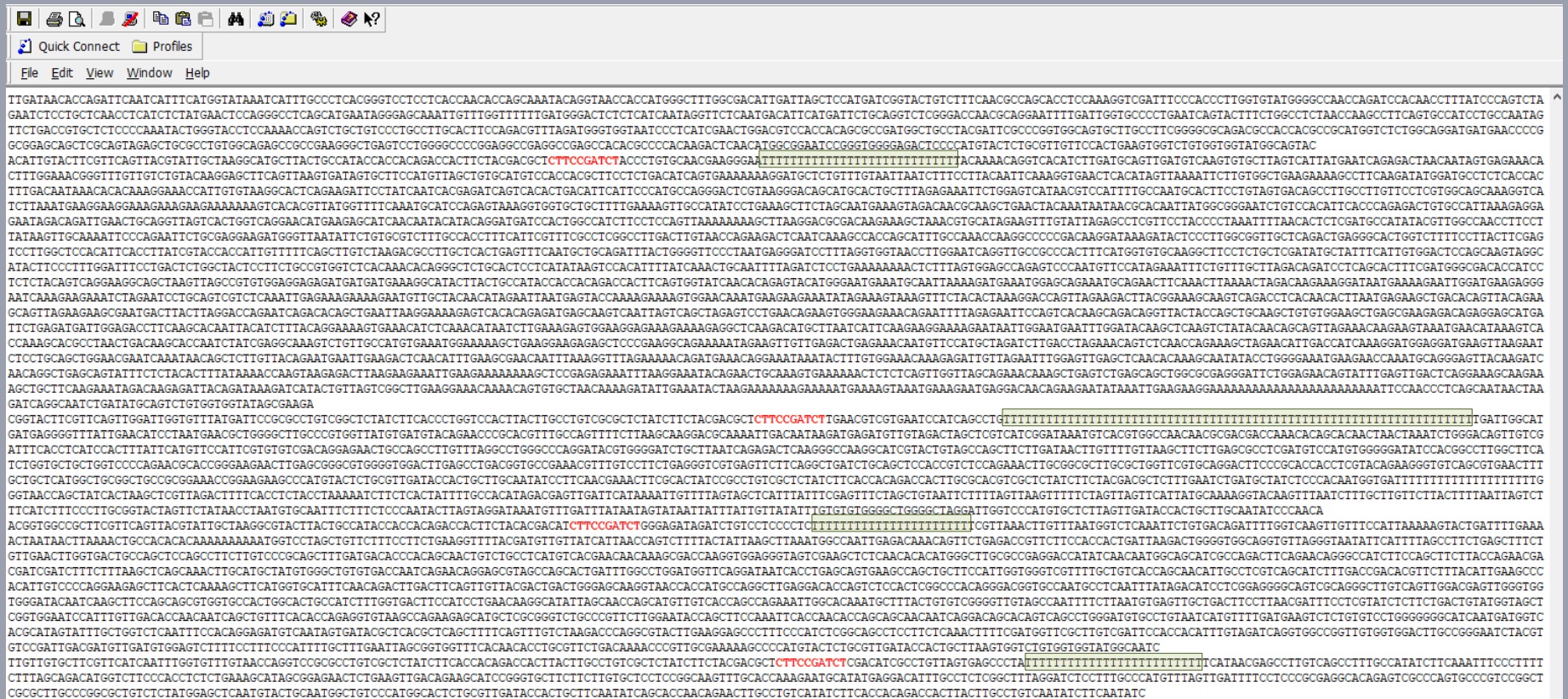
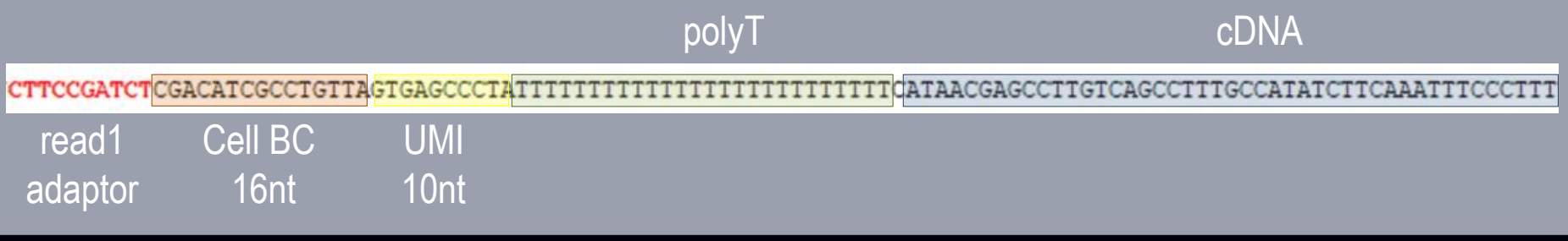


Plate of 96 B cells, Smart-seq2 modified protocol with Rolling Circle Amplification (RCA) and sequence with 1D R9.5 ONT flowcells. Good correlation with PacBio profiling, R2C2 can be easily adapted to any RNAseq library preparation protocol (10x, smart-seq2, drop-seq)

10x Genomics Nanopore reads



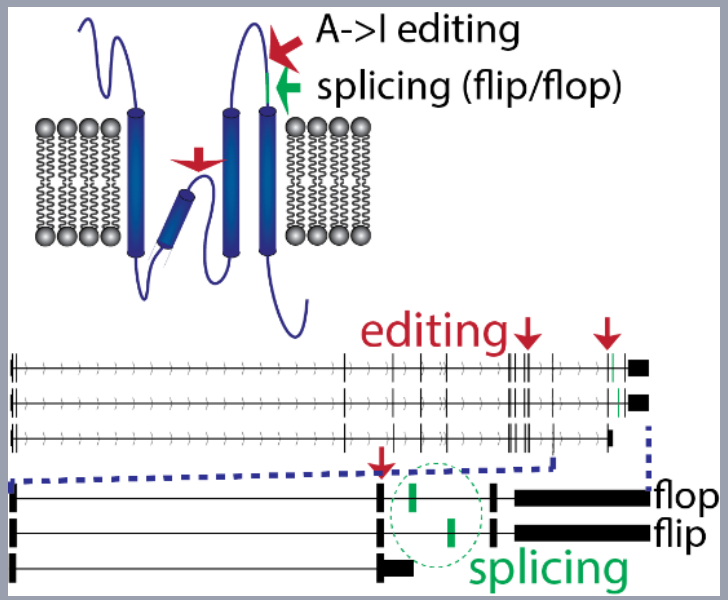


E18 mouse brain single cell transcriptome (Illumina and Nanopore profiling)

190 cells → Nanopore sequencing

951 cells → Targeted Nanopore sequencing

Targeted sequencing of the ionotropic glutamate (AMPA) receptor GRIA2



Targeted Gria2 Nanopore sequencing in 951 single cells