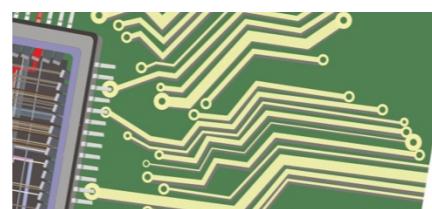


Single cell RNA-seq methods

Jellert Gaublomme

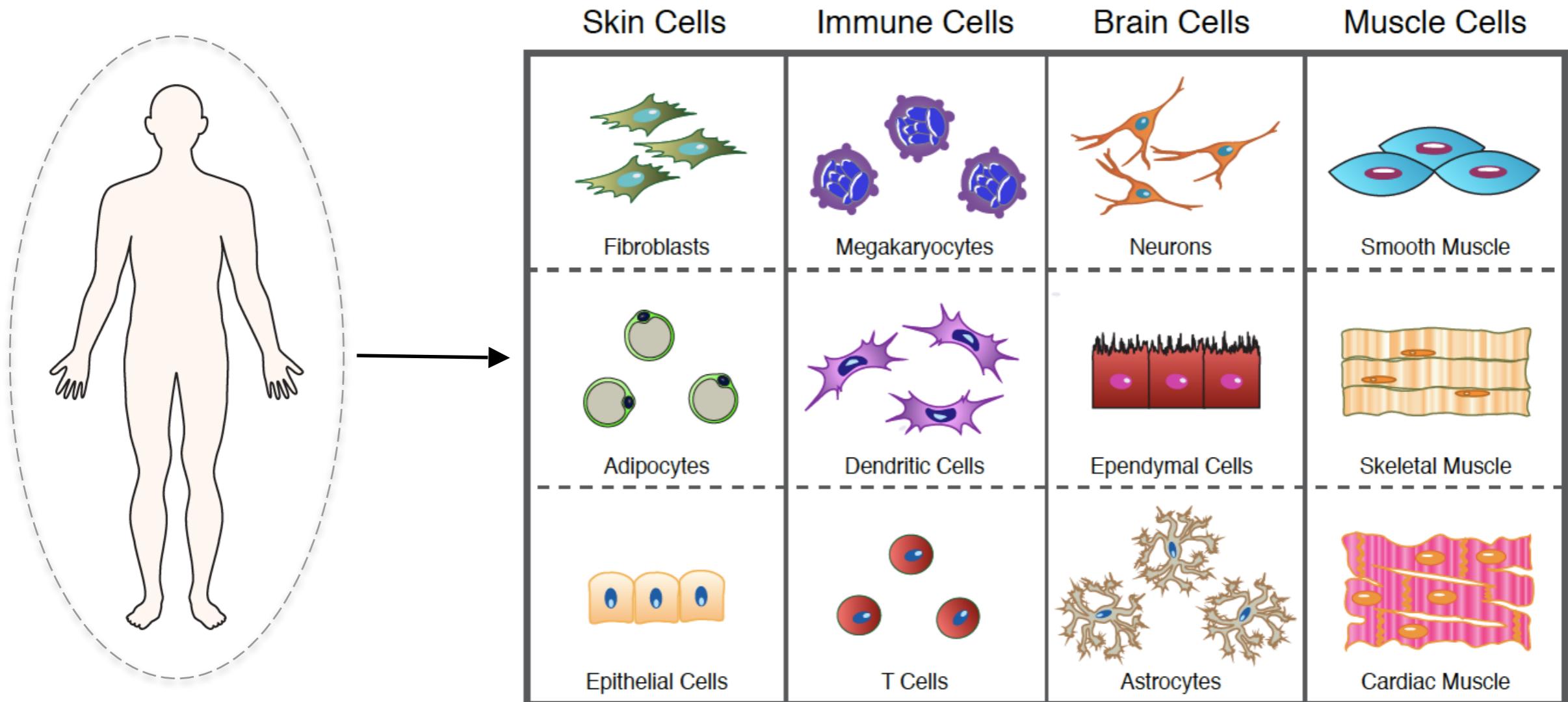
Klarman Cell Observatory
Broad Institute of MIT and Harvard

Slides by: Monika S. Kowalczyk, MD PhD



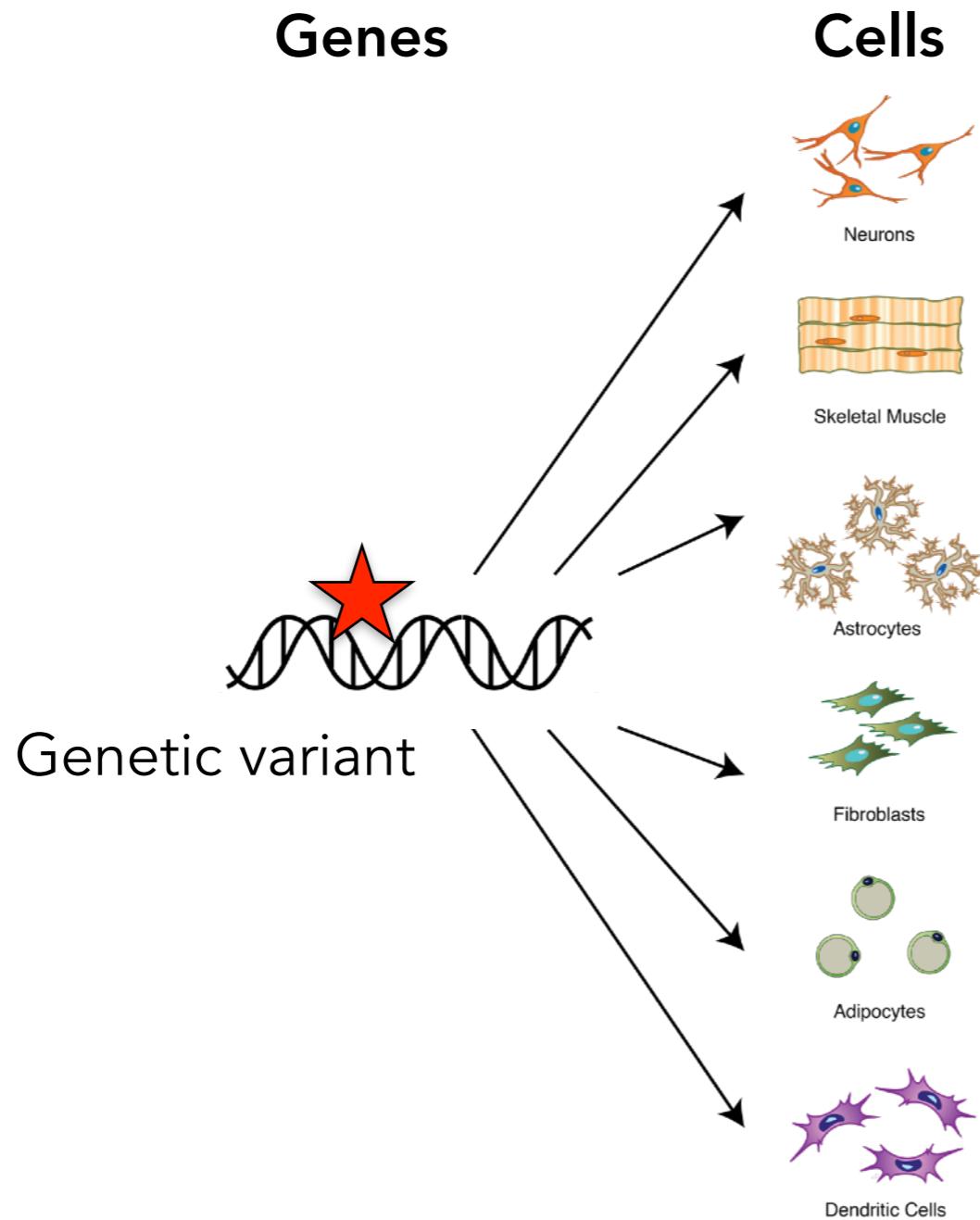
Computational
Genomics
Workshop
September 10 & 11, 2018

Cells are the basic unit of life



Cells are classified by structures, functions, location, and molecules

One genome, but many cells



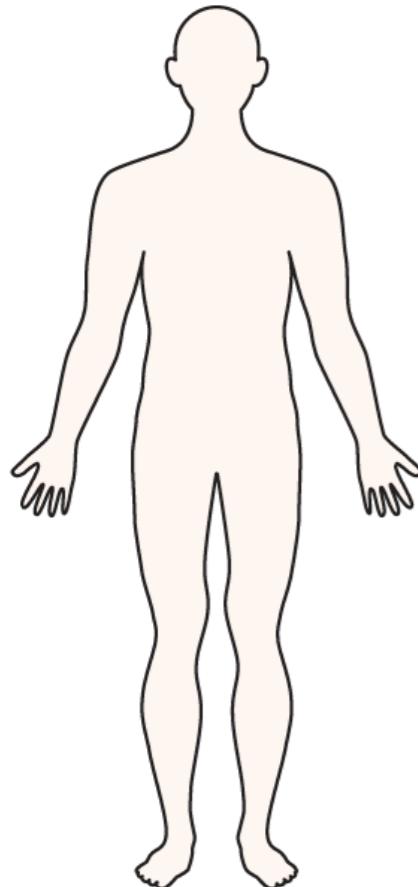
Knowing cells is essential to understand the genes that cause disease

Average may not represent population



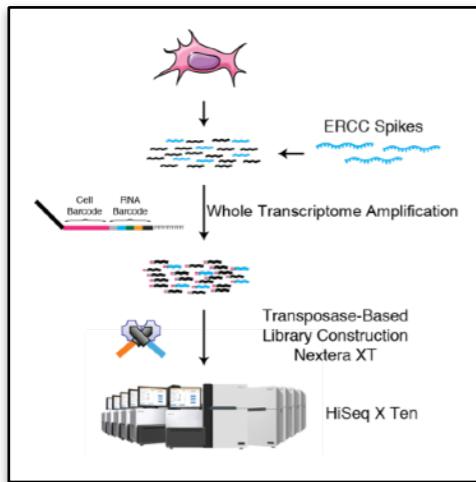
... although all are highly expressed on average

Problem: we do not know our cells

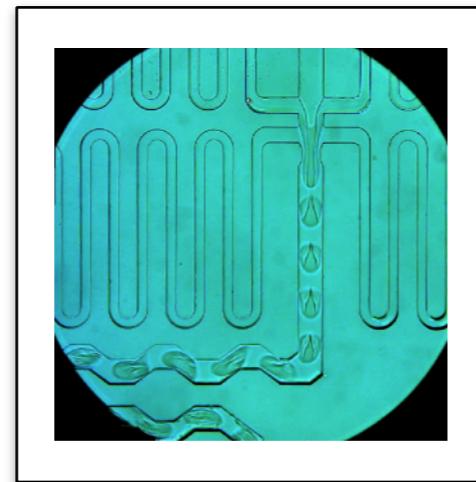


- **~30 trillion** cells
- Text book: ~300 'major' cell types?
- Science: ~100 sub-sub-types of retinal neurons

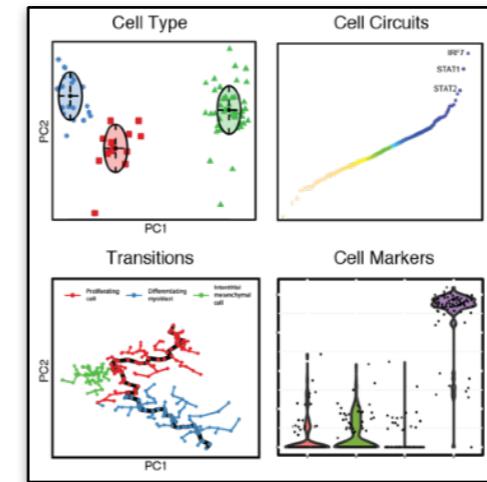
Technological advance: Single cell genomics



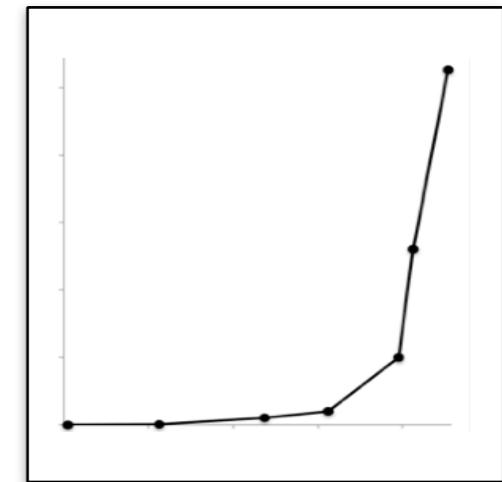
Core technology



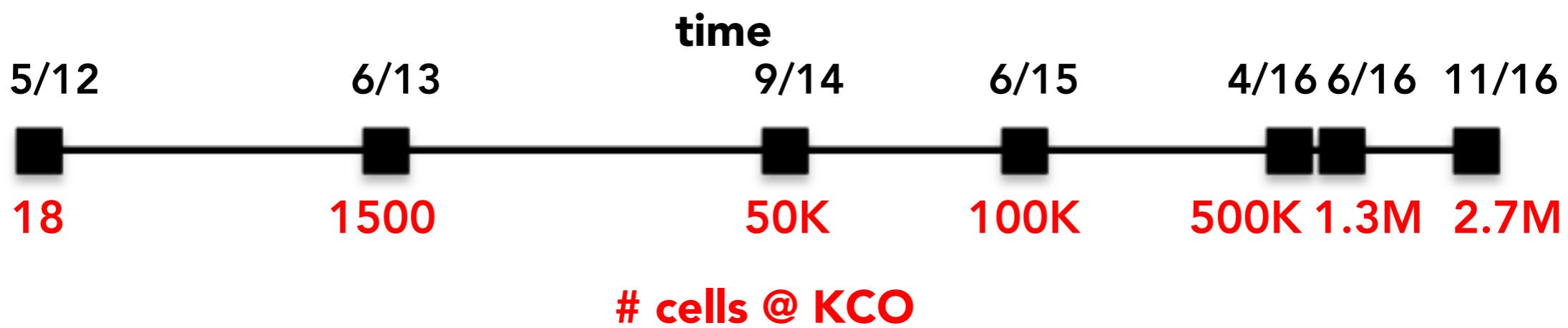
Sample prep



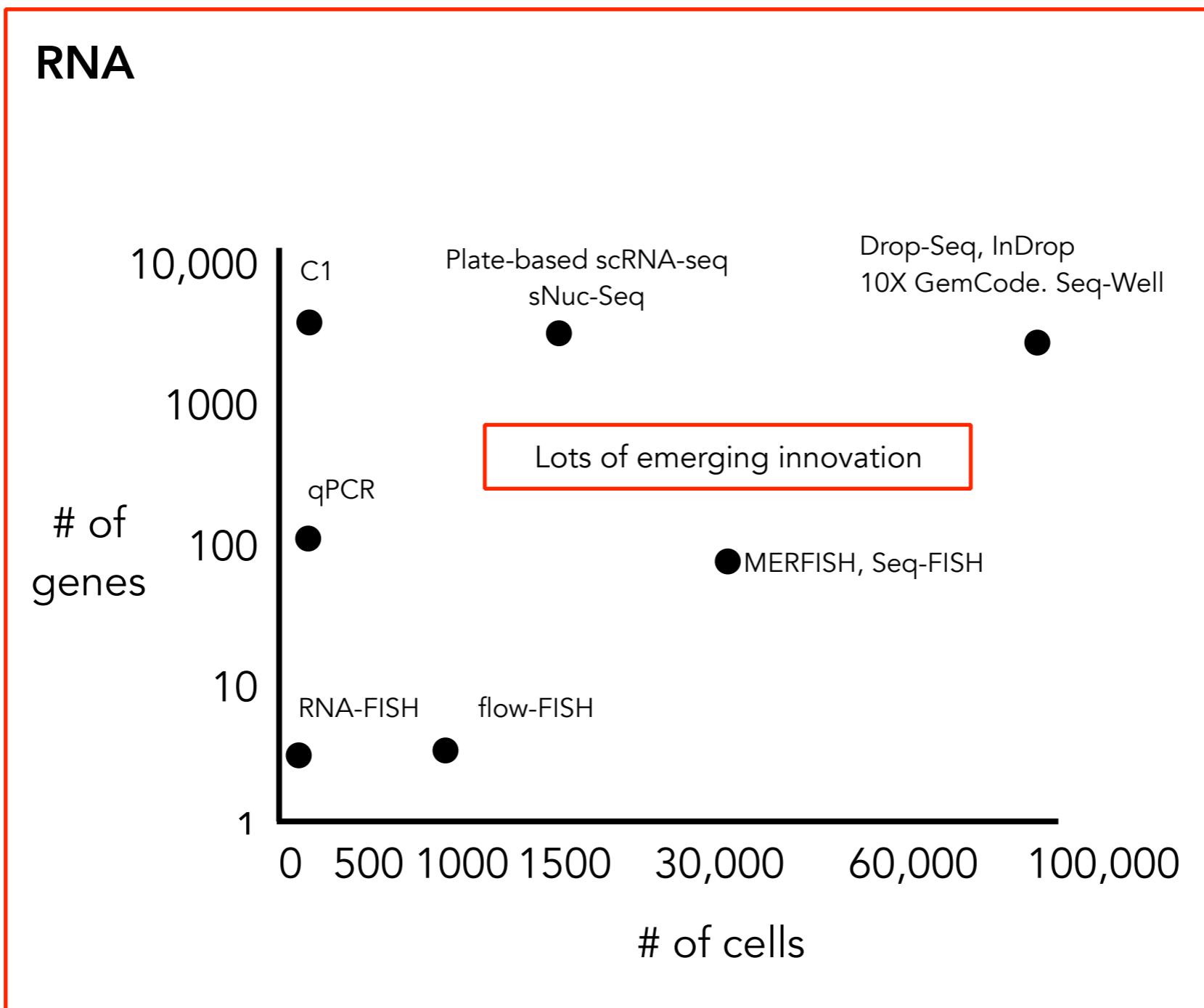
Computation



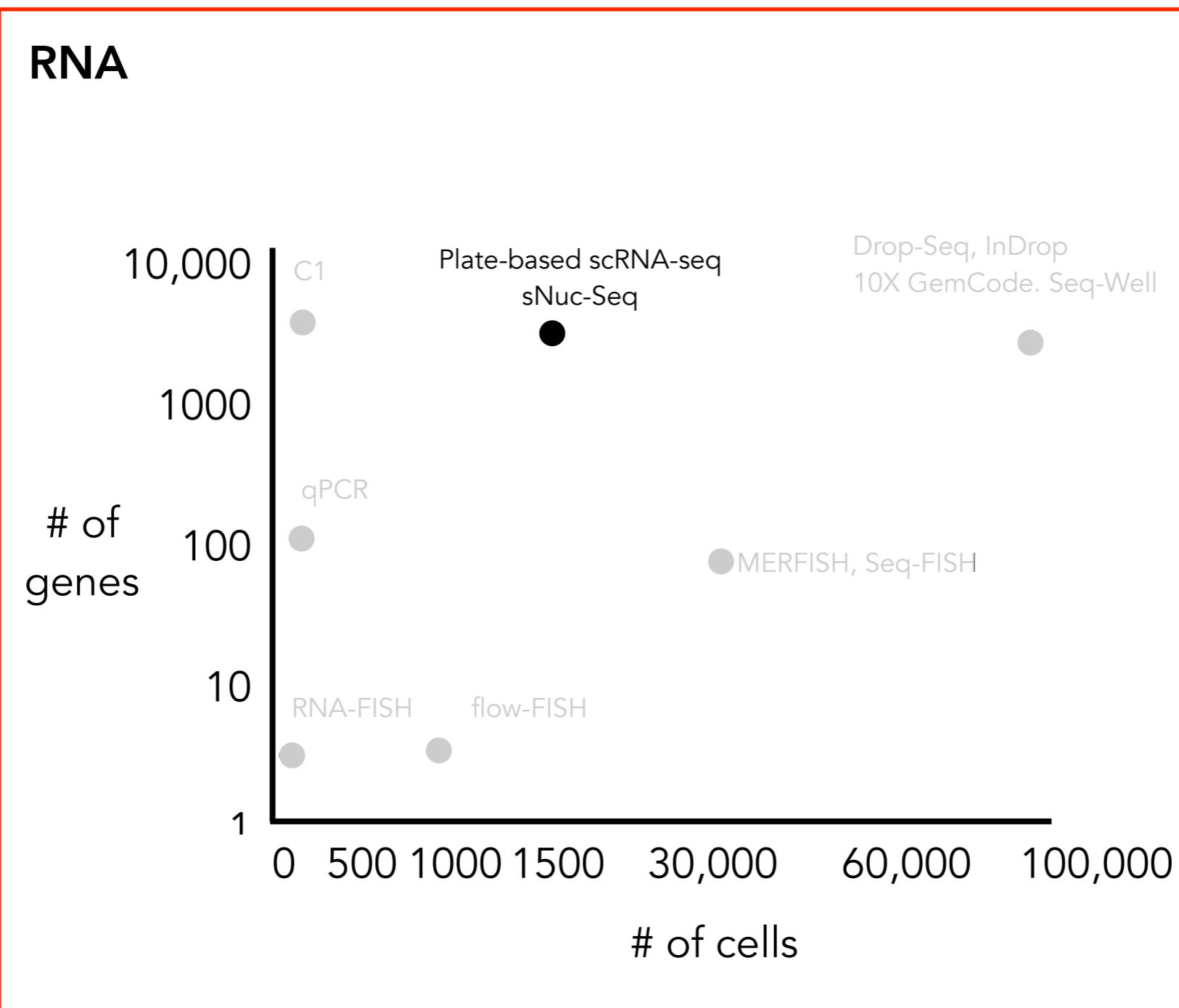
Massive scale



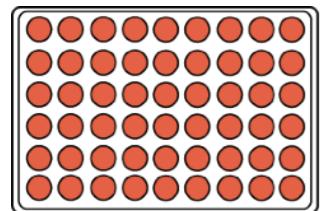
Tradeoff between scale & resolution



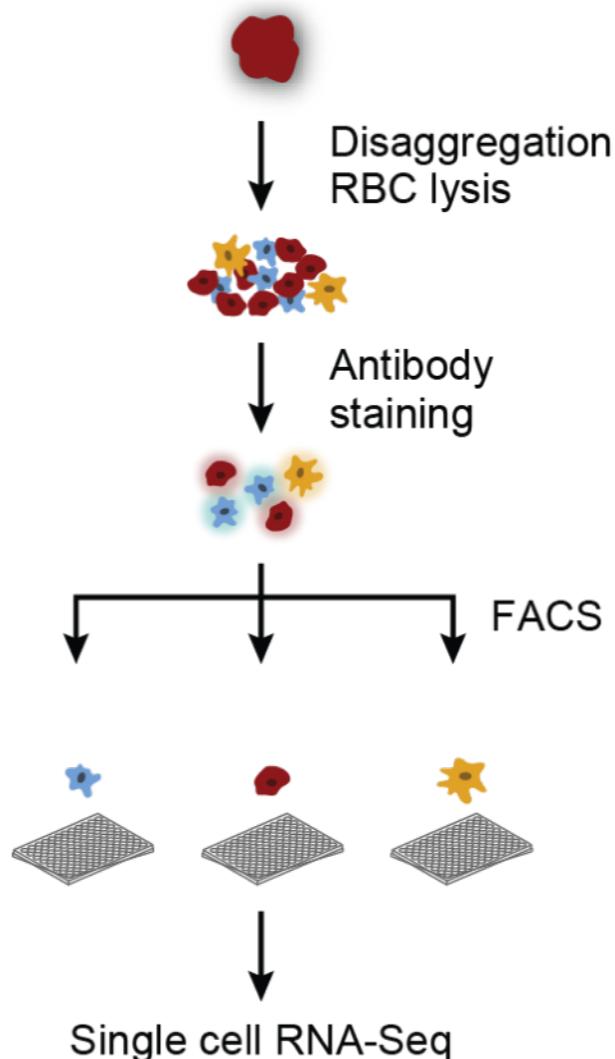
Tradeoff between scale & resolution



Single-cell RNA-seq pipeline



Plates



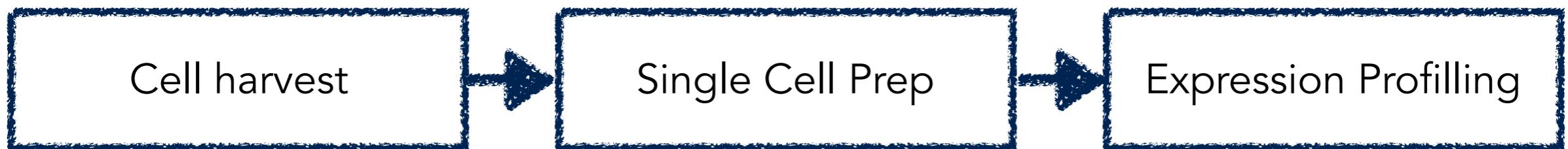
sNuc-seq

you can now process nuclei!

Habib et al *Science* 2016

up to 1,000 cells

Single-cell RNA-seq pipeline



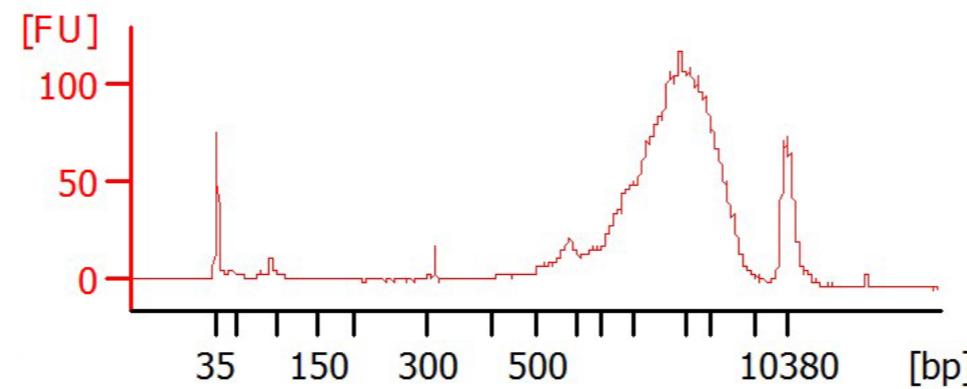
- reverse transcription
- whole transcriptome amplification

Expression Profiling

- library construction
- sequencing

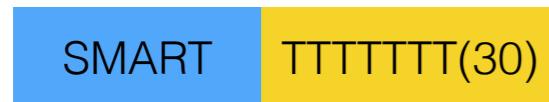


Single-cell RNA-seq pipeline

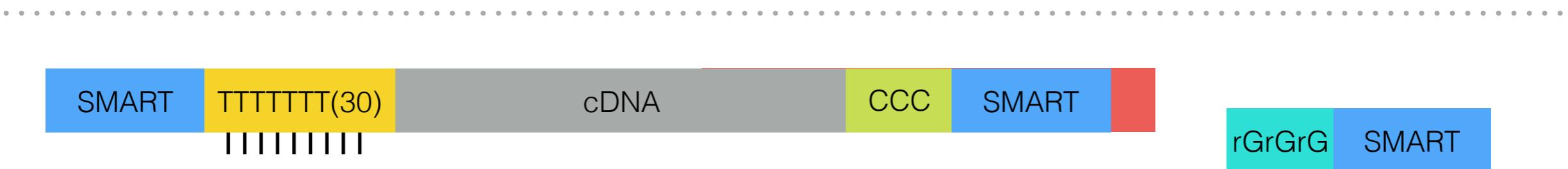


Whole transcriptome amplification product

primers



mRNA



Single-cell RNA-seq pipeline

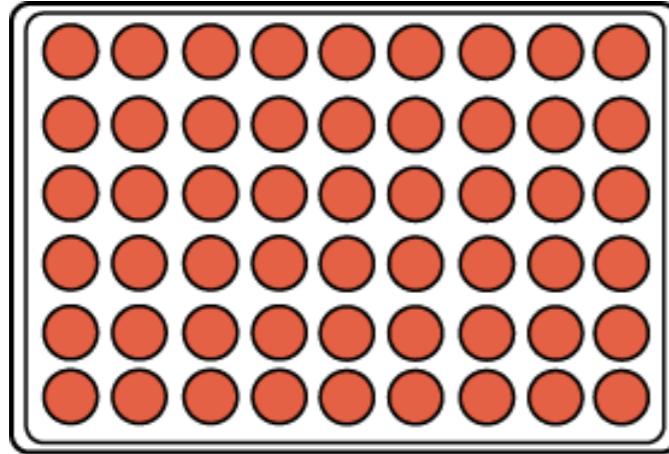
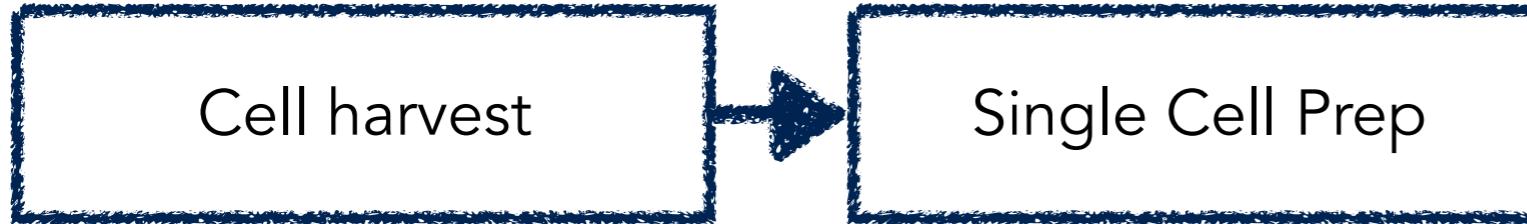
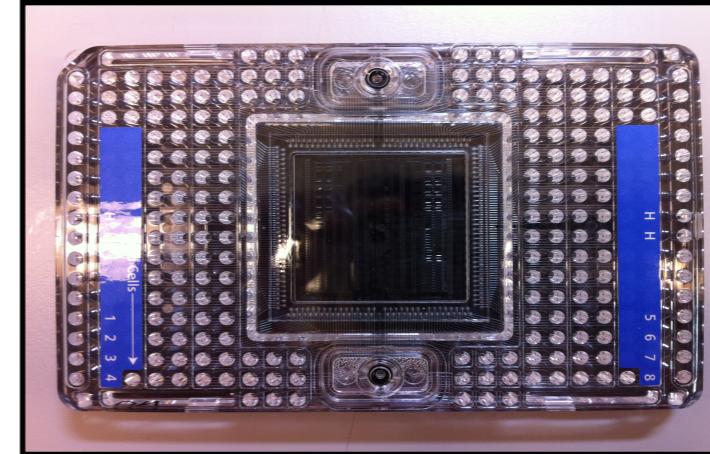


Plate approach (96 or 384)

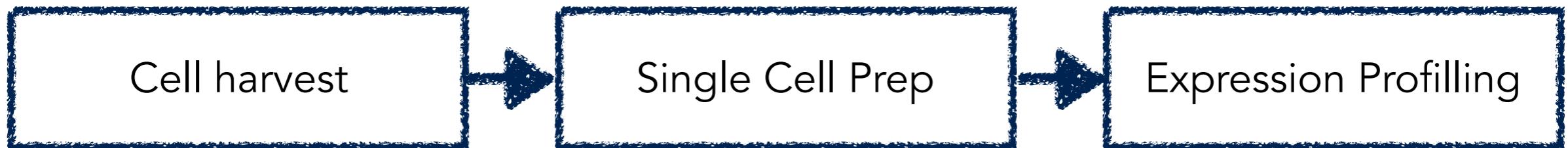
- archive cells
- select specific cell
- index sorting (protein levels)
- labor intensive
- costly (\$10 per cell)



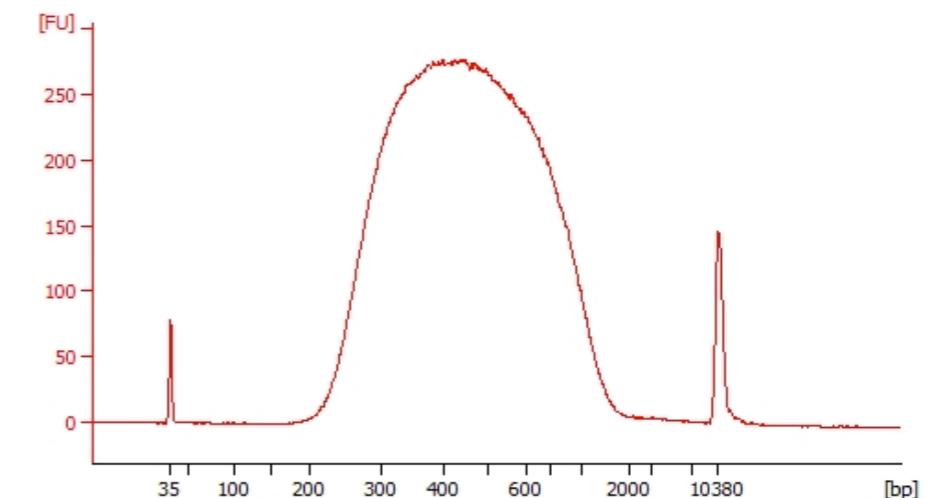
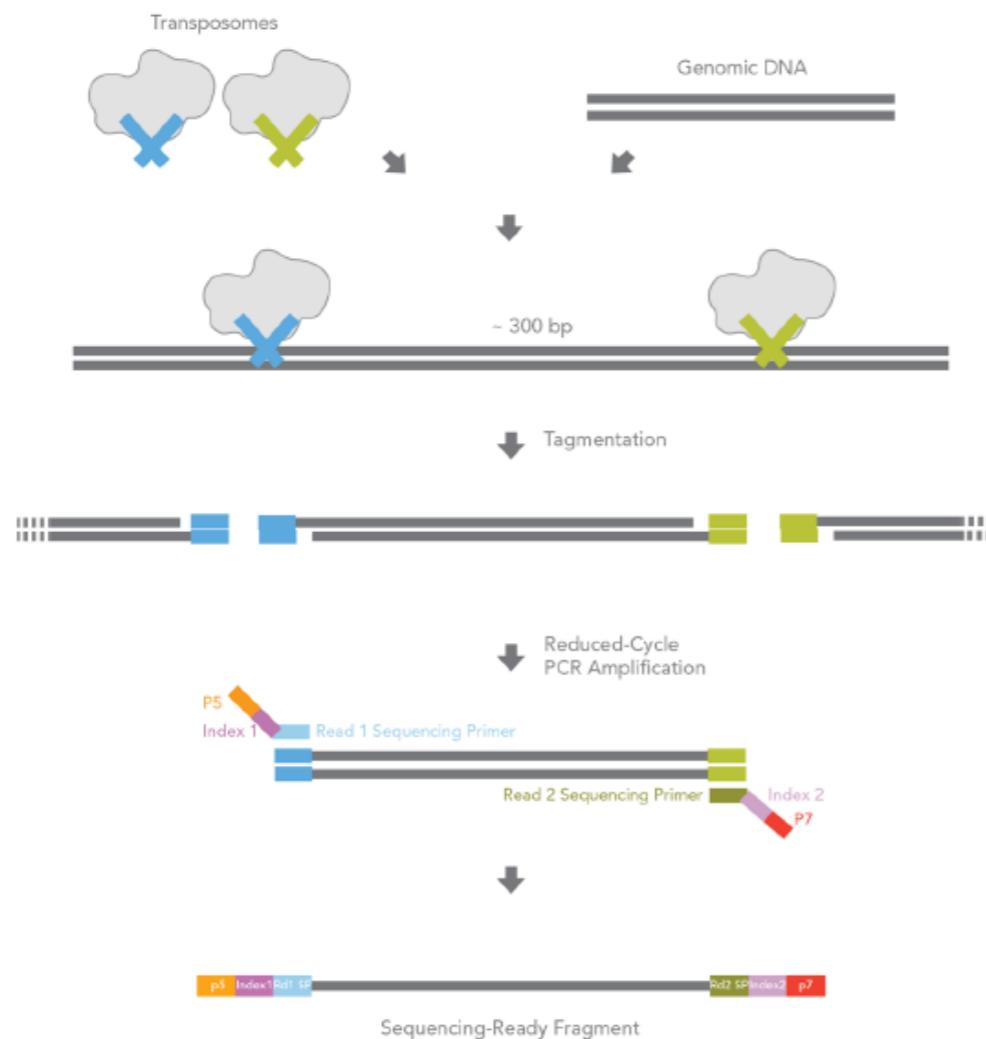
C1 (96 or 800)

- easily profile up to 800 cells
- minimum cell input
- nano-scale reaction volume
- cannot select specific cells

Single-cell RNA-seq pipeline



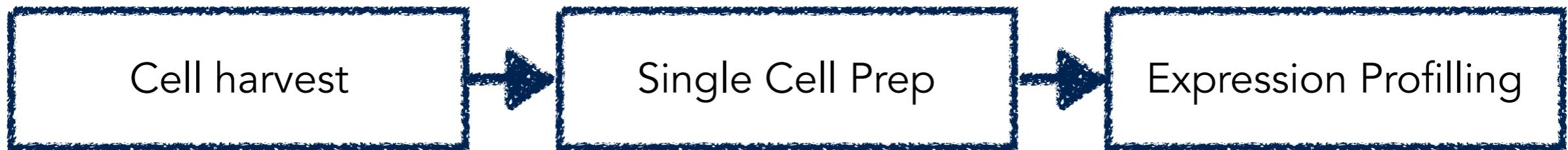
NexTera XT



- pool libraries
- SPRI clean up
- quantify

384 cells into libraries in less than 4 hours

Single-cell RNA-seq pipeline



Miseq



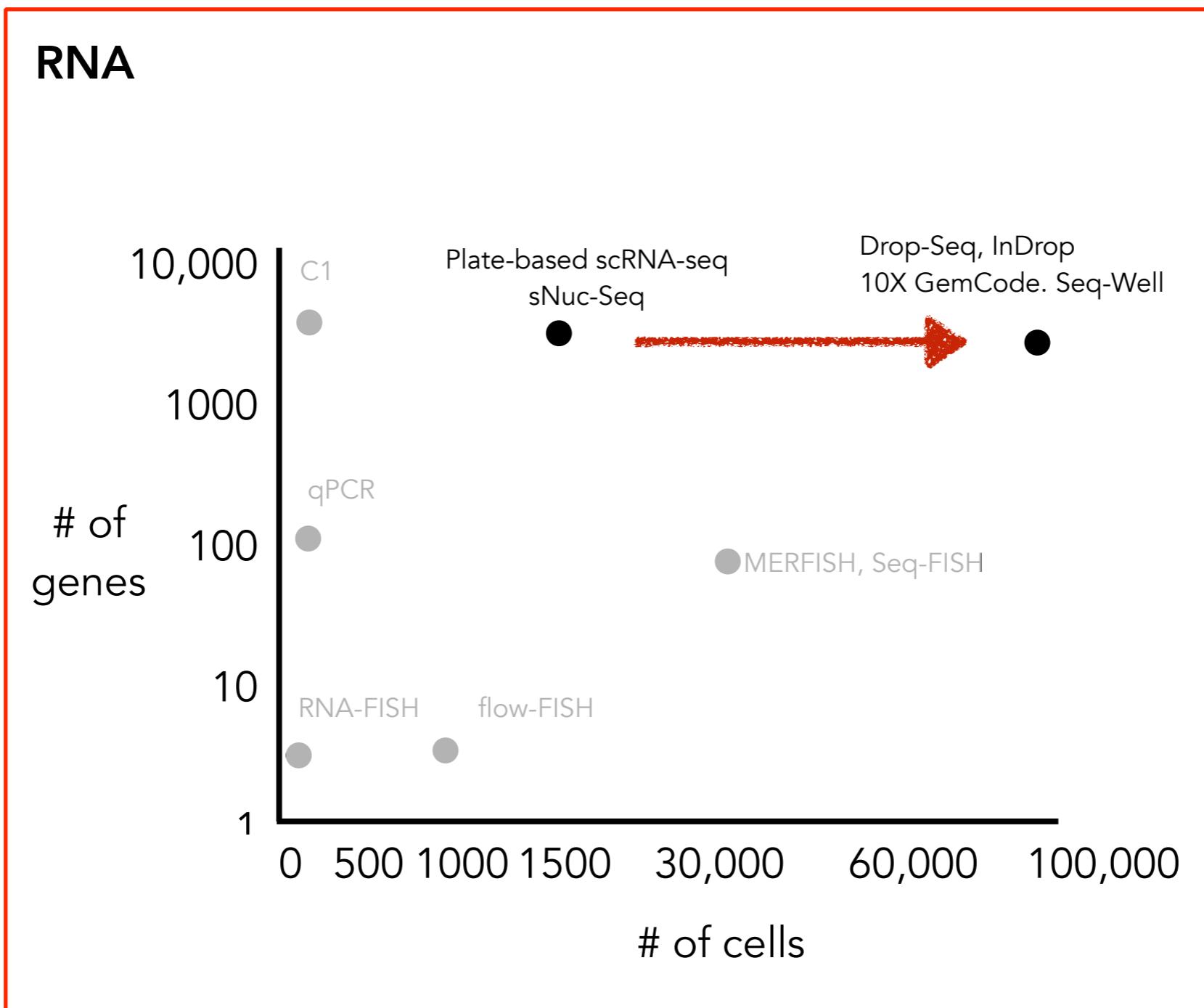
~ 20M reads total
~ up to 96 cells per run

Nextseq

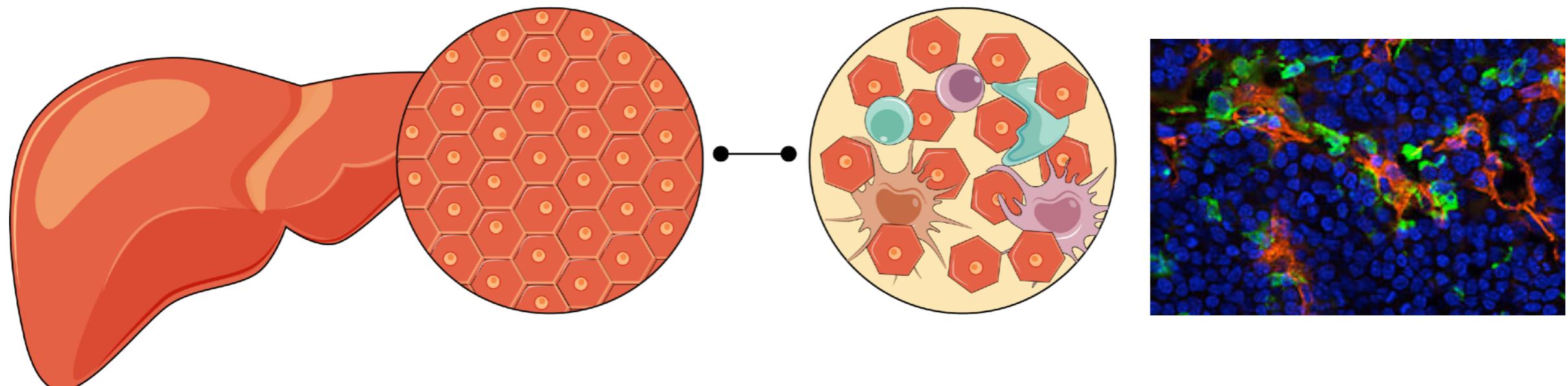


~ 500M reads total
~ up to 1000 cells per run

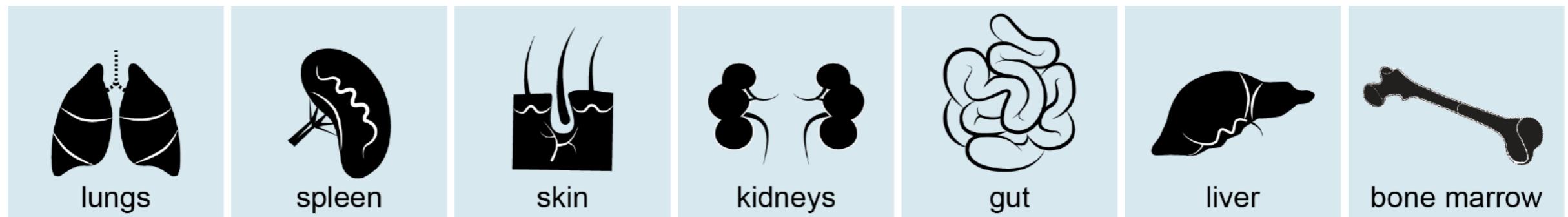
Tradeoff between scale & resolution



scRNA-seq at scale



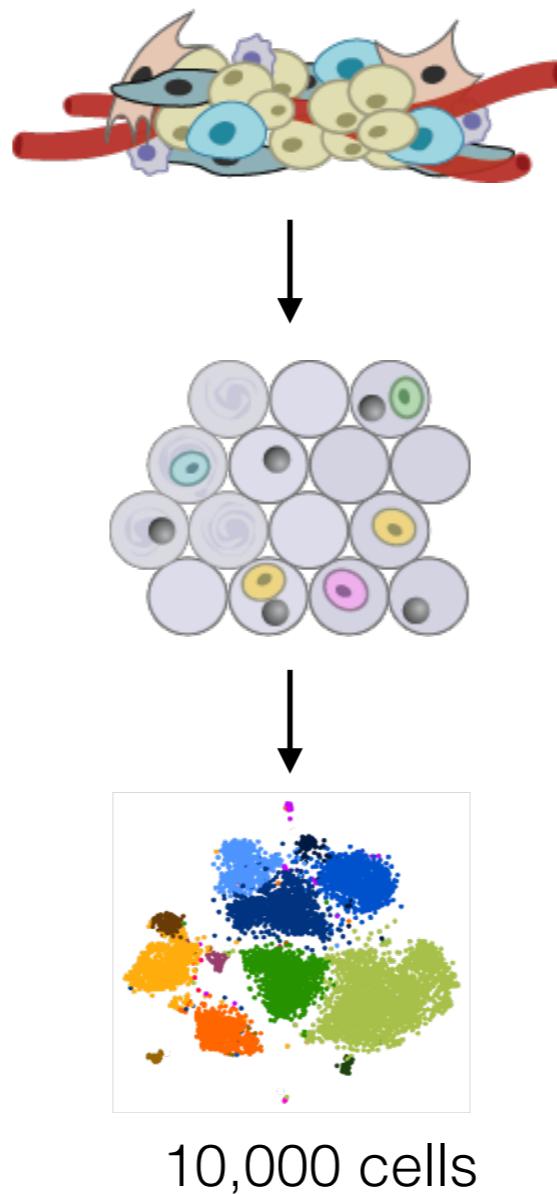
tissue specific cell types
immune cells



Can we identify cell types and cell states?

scRNA-seq at scale

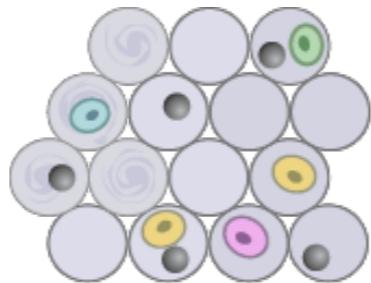
Droplet based approaches



going from hundreds to tens of thousands of cells

scRNA-seq at scale

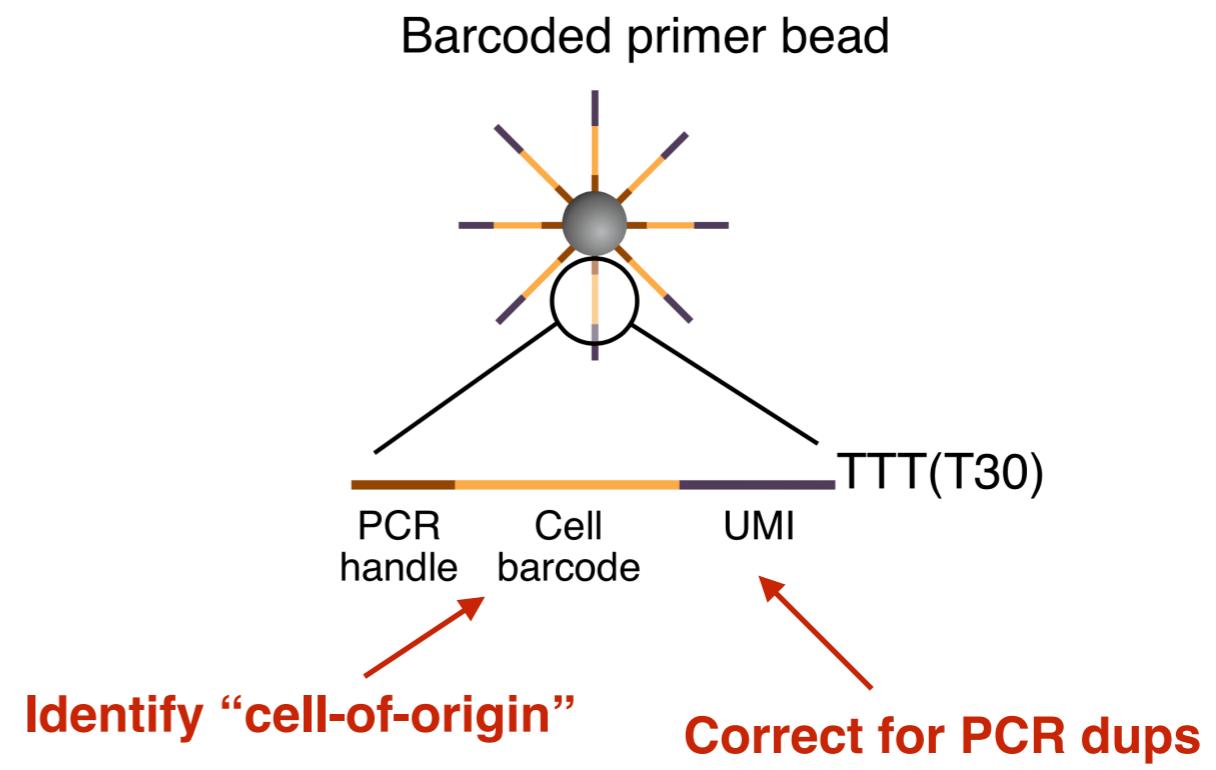
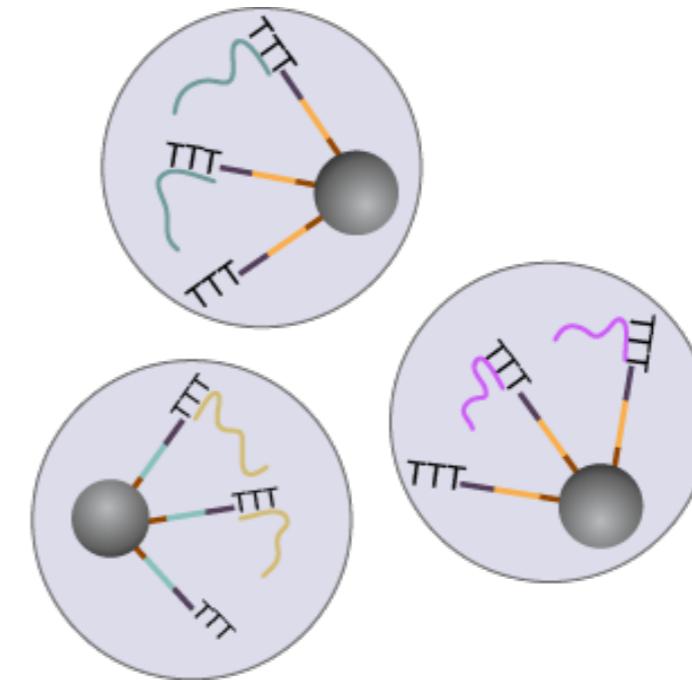
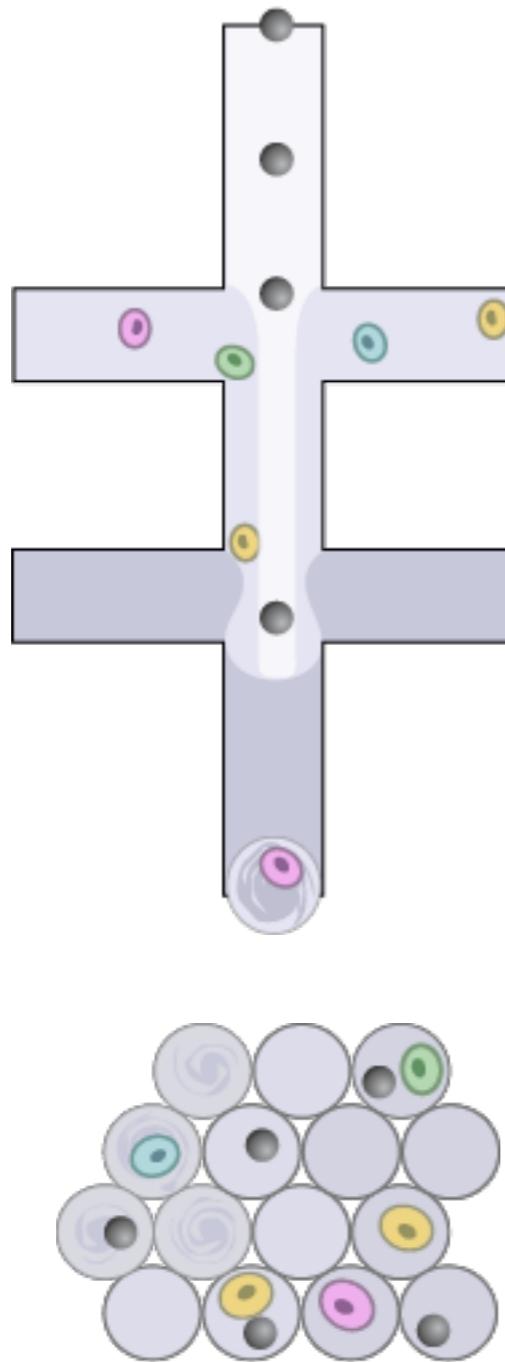
Droplet based approaches



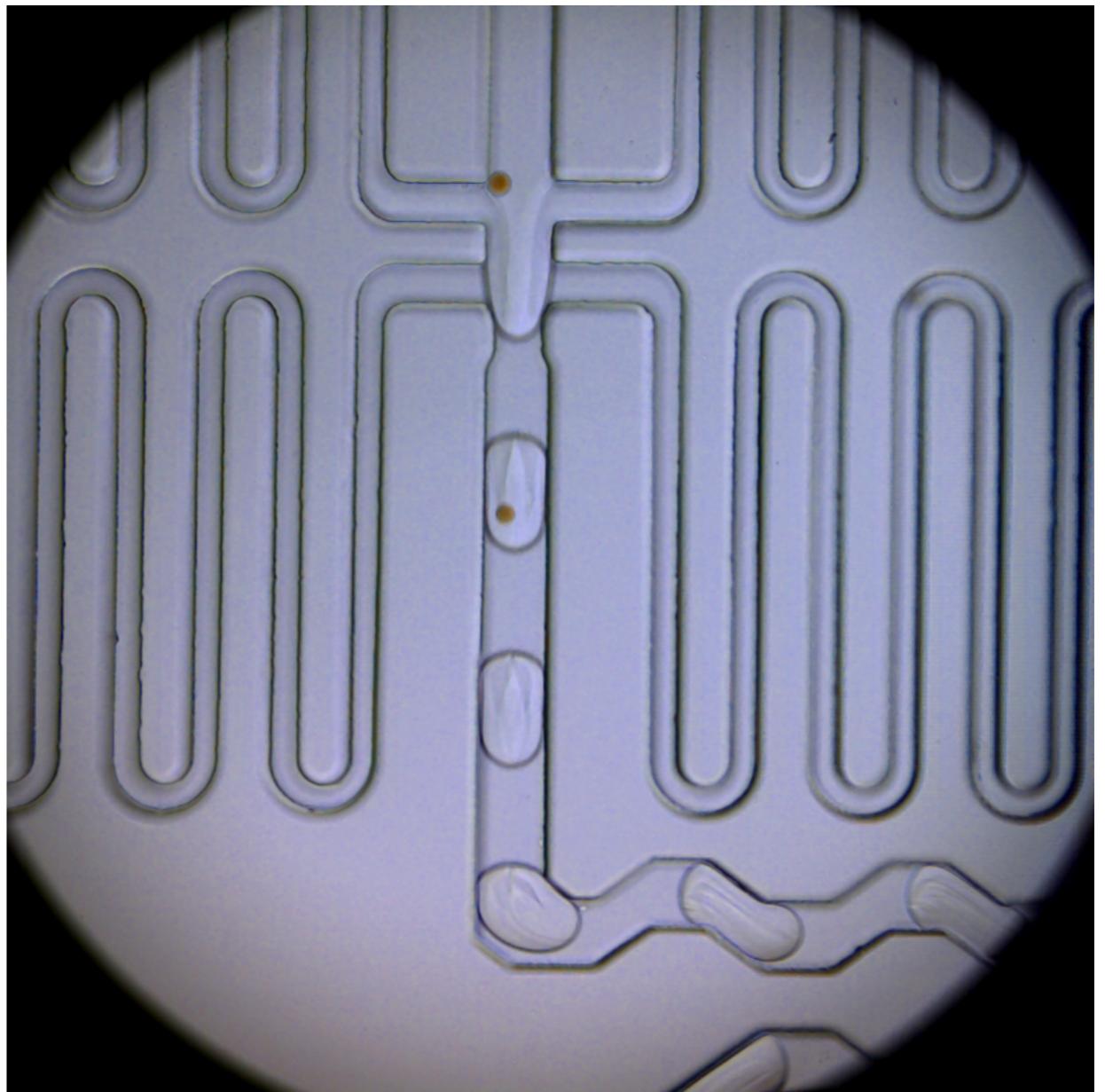
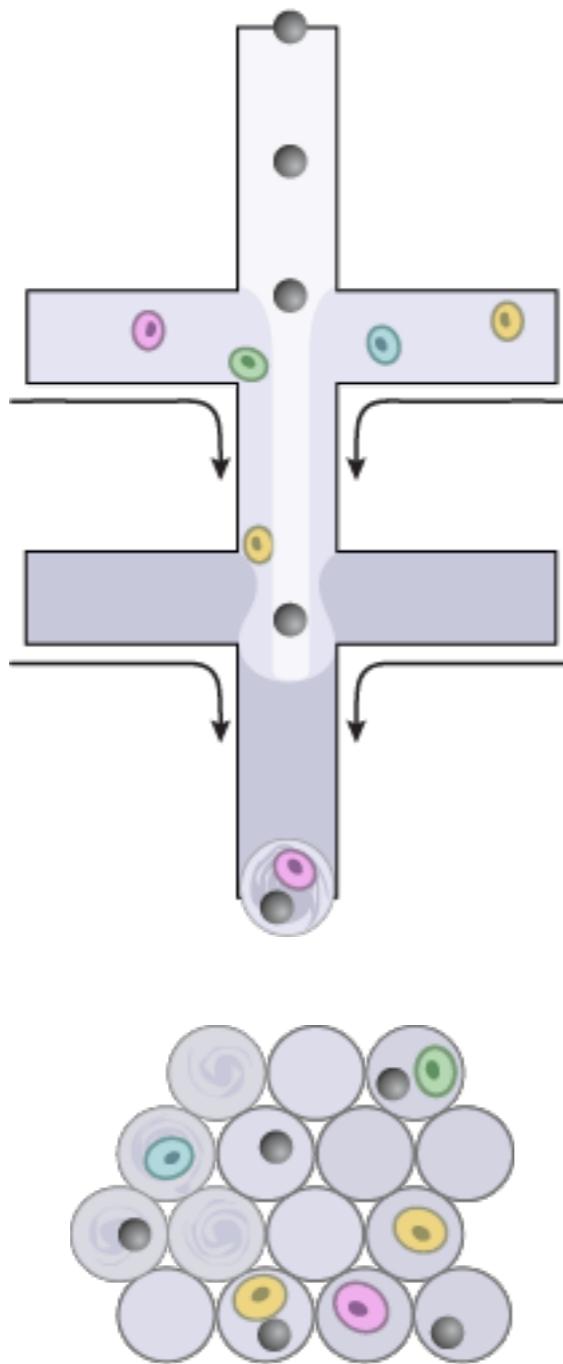
1. Drop-seq
2. InDrop
3. 10X Genomics

Massively parallel approaches are 3'end!

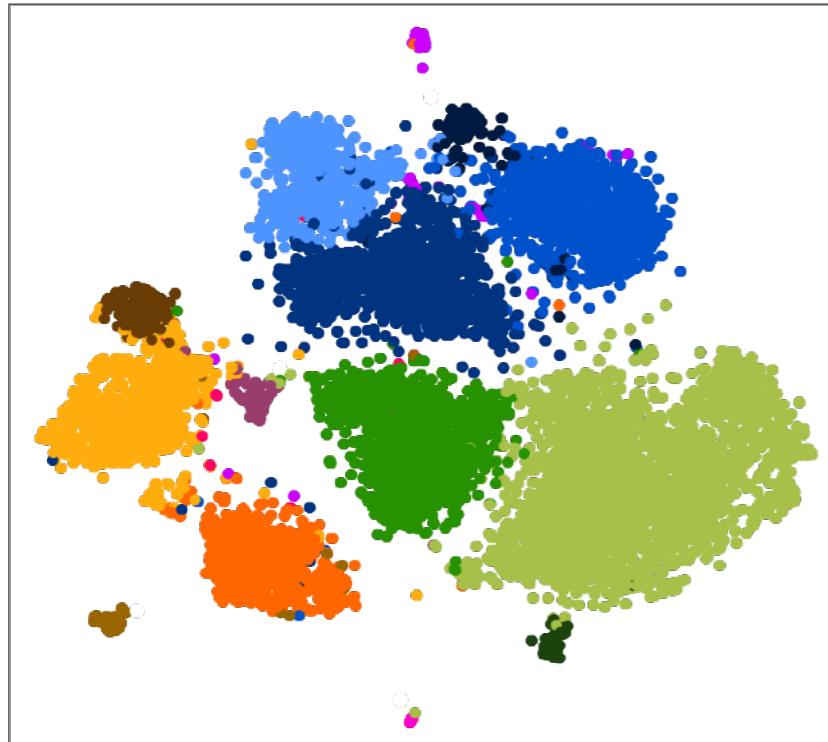
Drop-Seq: massively parallel scRNA-seq



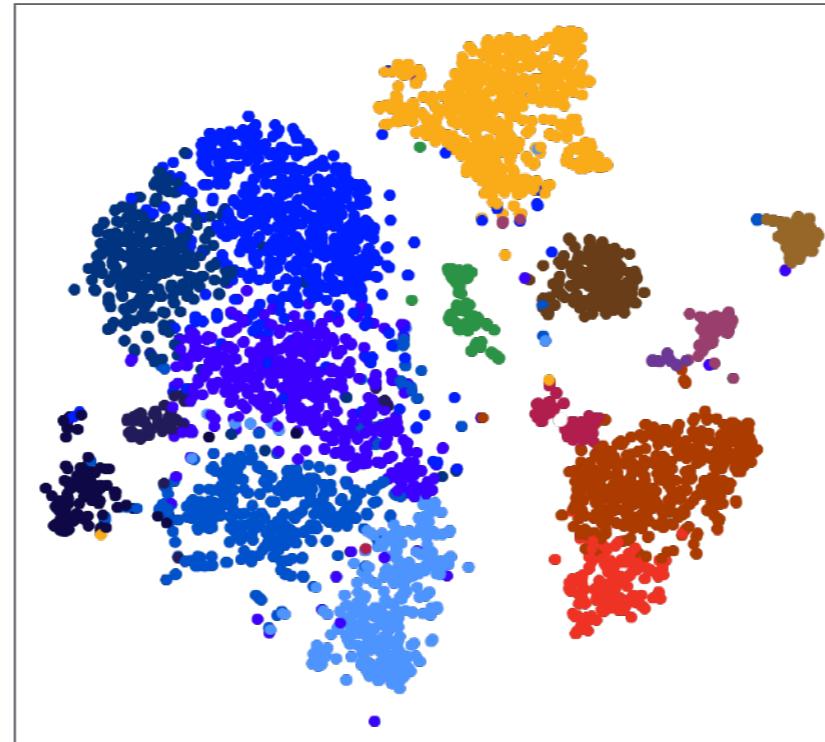
Drop-Seq: massively parallel scRNA-seq



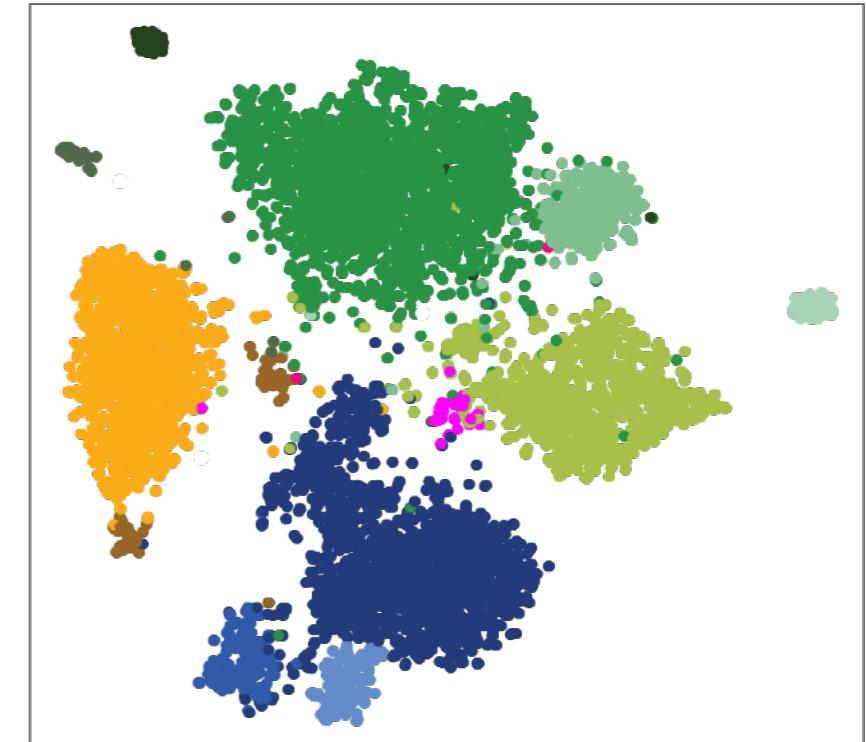
Immune cells in peripheral tissues



10,000 cells



10,000 cells

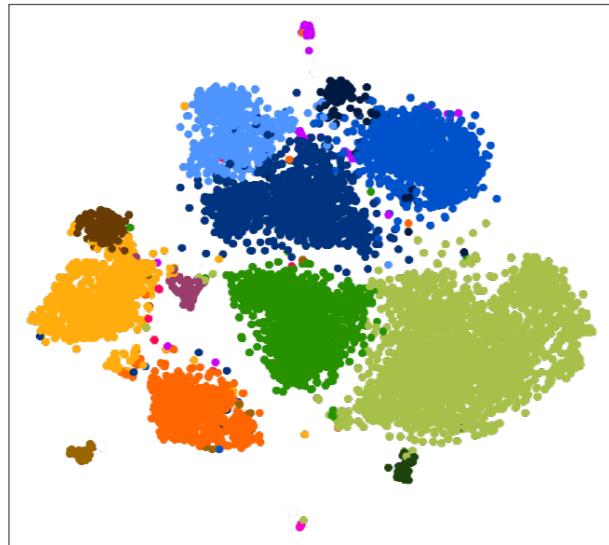


10,000 cells

you can now use fixed cells!

Alles, Praktiknjo, Karaiskos et al *Biorxiv* 2017

Massively parallel scRNA-seq approaches



Dropseq

- profile up to 2000 cells per 15 min
- requires 100,000 cells as input
- labor intensive
- cheap (20 cents per cell)



10X GENOMICS®

10X Genomics

- parallel channels
- profile up to 4000 cells per channel
- requires 7,000 cells as input
- 8 minutes for droplet generation
- from cells to library in one day
- upfront costs (kit ~25K)
- 40 cents per cell

An evolving technology landscape

RNA

- Full length
(mRNA, total RNA)
- 5' and 3' end counting

'multi-omics'

- DNA+RNA (G+T)
- RNA+protein (T+P)
- Epigenome + RNA

Spatial

- Multiplex FISH (Seq-FISH, MERFISH)
- In situ* RNA-Seq (e.g., FISSEQ)

RNA + protein

CITE-seq (Stoeckius et al Nature Methods 2017)

REAP-seq (Peterson et al Nature Biotechnology 2017)

Split and pool barcoding (non-microfluidics)

SPLiT-seq (Rosenberg et al BioRxiv 2017)

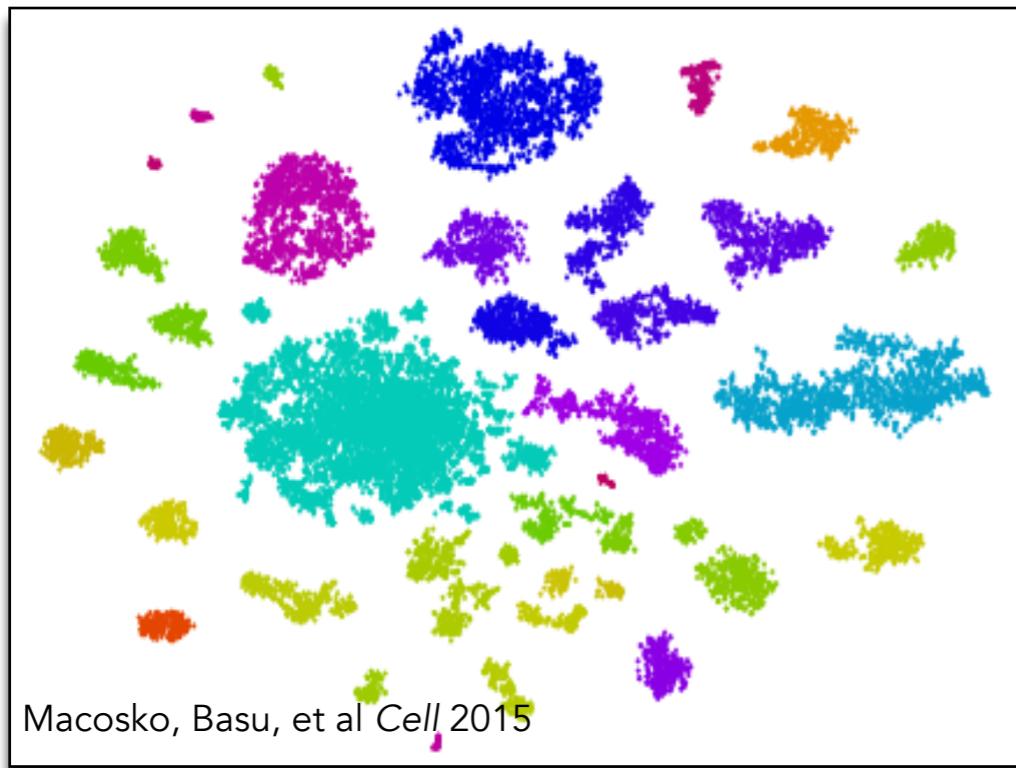
SCI-seq (Cao et al Science 2017)

What can we learn from single cells?

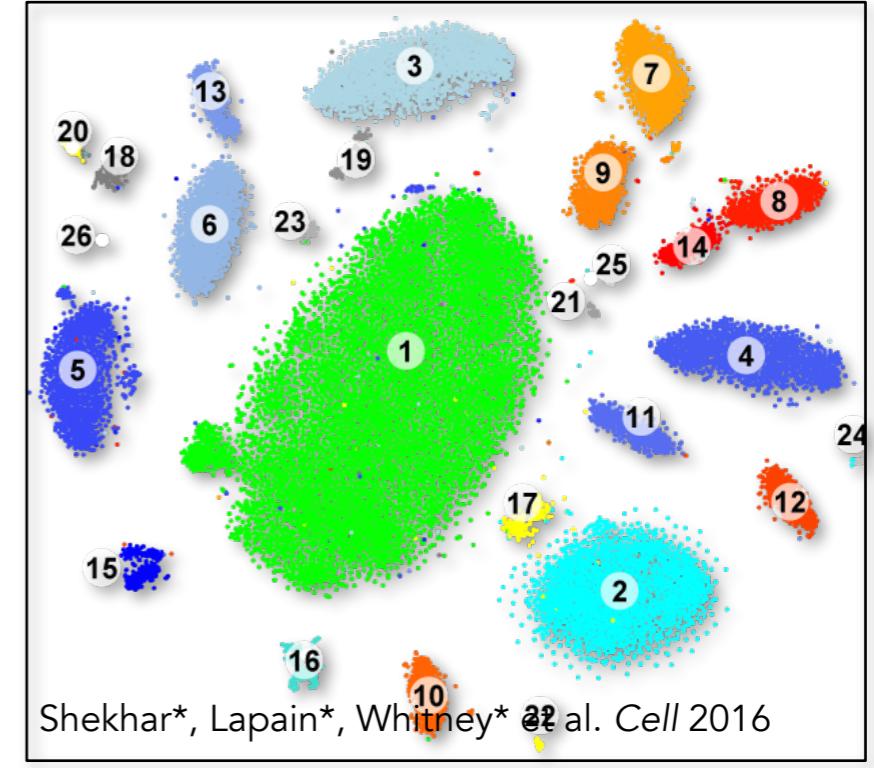
- ✓ **Taxonomy:** cell types
- ✓ **Histology:** tissue structure
- ✓ **Developmental biology:** cell fate and cell lineage maps
- ✓ **Physiology:** cycles, transient responses, plastic states
- ✓ **Pathology:** disease cells in their tissue ecosystem
- ✓ **Molecular mechanisms:** intra- and inter-cellular circuits

Taxonomy: Cell types

Whole retina (48,808 cells)

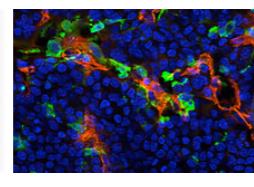
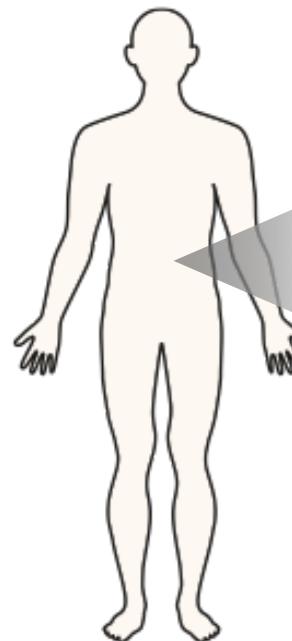


Bipolar cells (28,000)

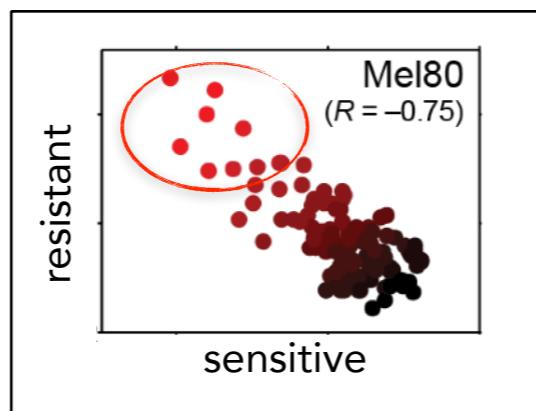
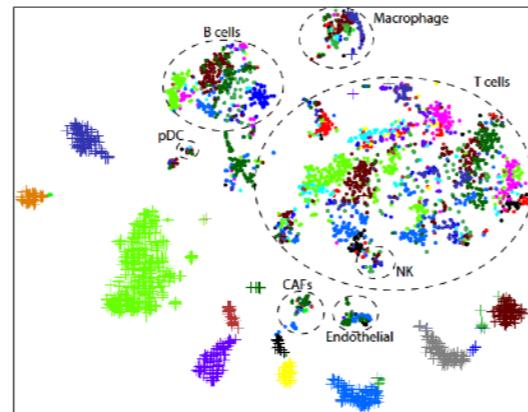


- **Goal:** Molecular definition of all cell types; harmonization with other taxonomies
- **Cell type:** a region or probability distribution in full space or projection
- **Intuitively compelling, surprisingly elusive:** new data-driven definitions?
- **Challenges:** Sampling, computational scalability, over-categorization, validation

Disease: the cellular ecosystem

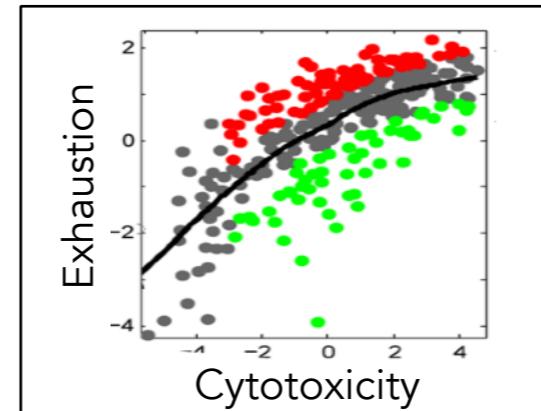
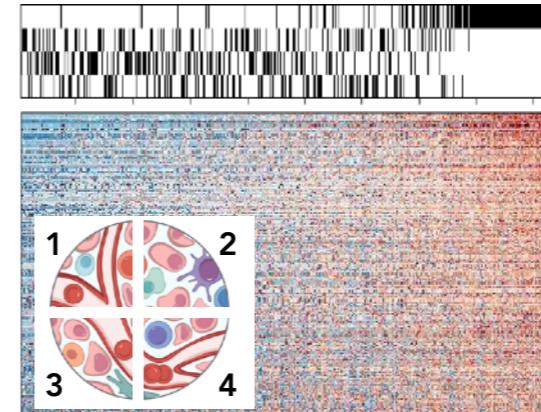


1. Taxonomy



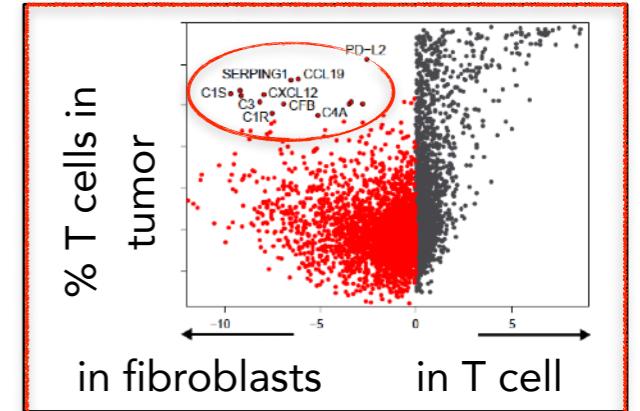
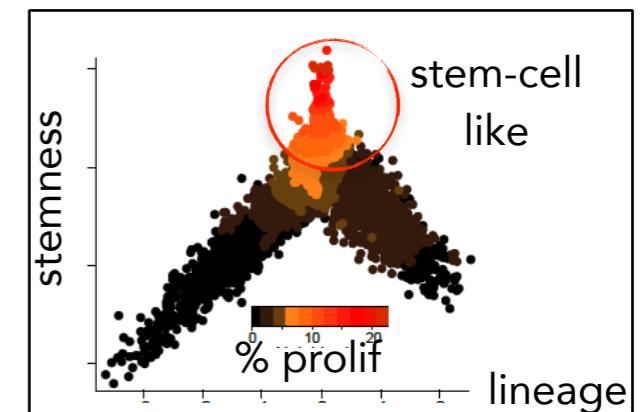
4a. Physiology

2. Histology



4b. Physiology

3. Dev. biology

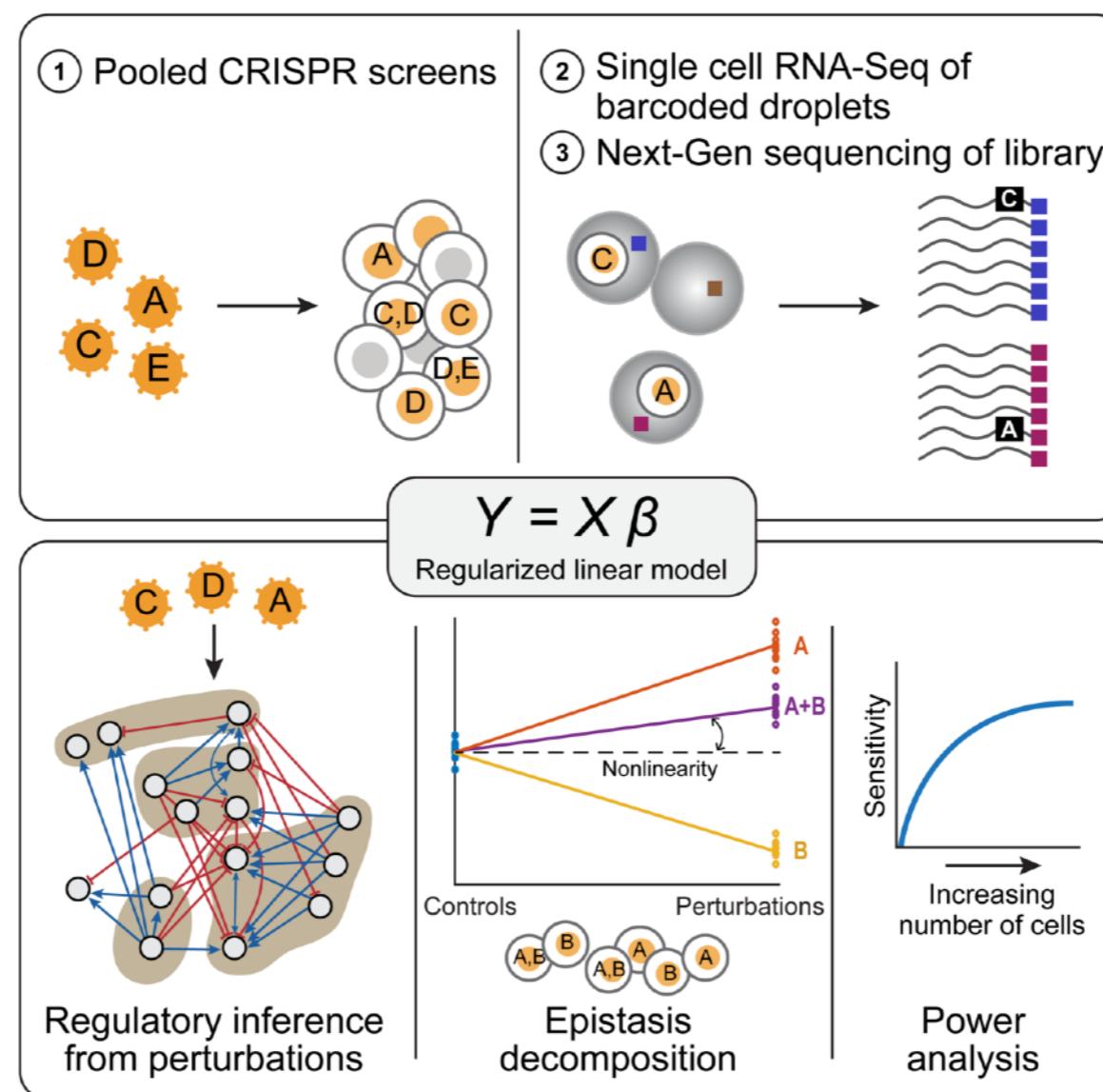


5. Cell interactions

- **Goal:** Changes in cell proportions, compositions and intrinsic states, and the inter- and intra-cellular circuits that govern them
- High-res comparison across patients, systems: 'bulk' deconvolution, model 'faithfulness'
- **Challenge:** 4D analysis: molecular, spatial, and temporal (with genetics)

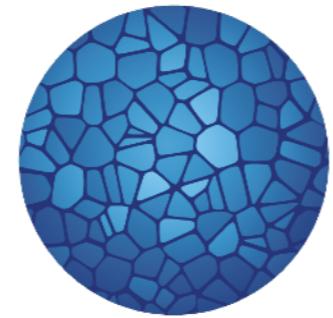
From cell atlas to molecular mechanisms

Example: Perturb-Seq



Building the HCA community

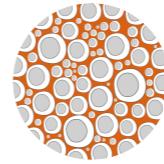
Biological networks



Technology forum



HUMAN
CELL
ATLAS



Data platform

portals.broadinstitute.org/single_cell



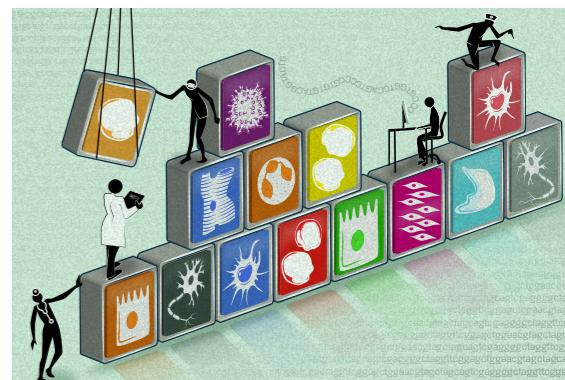
Analysis garden

Steering group, white papers, meetings
www.humancellatlas.org

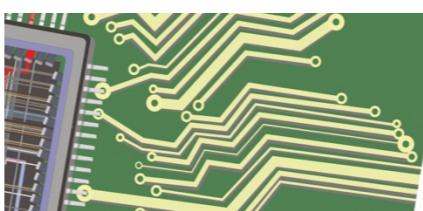
Aviv Regev with Sarah Teichmann (EBI/Sanger)

Organizing Team: Mike Stubbington, Angela Macharia, Orit Rosen, Jane Lee

Organizing Committee: Mike Stratton, Eric Lander, Jay Shin, Alex van Oudenaarden, Sten Linnarsson, Ehud Shapiro, Peter Campbell, Cori Bargmann, Steve Quake, Jonathan Weissman, Garry Nolan, Arnold Kriegstein, Christophe Benoist, Nir Hacohen, Ramnik Xavier, Hans Clevers, Dana Pe'er, Ido Amit, Chris Ponting, Piero Carninci, Barbara Wold



Thanks to Regev group and Klarman Cell Observatory



Computational
Genomics
Workshop
September 10 & 11, 2018

Xian Adiconis, Asma Bankapur, Inbal Benhar, Moshe Biton, Jason Buenrostro, Tyler Burks, Leah Caplan, Mary Carmichael, Jenny Chen, Brian Cleary, Ofir Cohen, Le Cong, Julianna Coraccio, Michael Cuoco, Inbal Davidi, Carl de Boer, Danielle Dionne, Atry Dixit, Eugene Drokhlyansky, Daneyal Farouq, Jeffrey Farrell, Leslie Gaffney, Jellert Gaublomme, Jonathan Gootenberg, Brian Haas, Adam Haber, Naomi Habib, Rebecca Herbst, Cindy Hession, Eran Hodis, Matan Hofree, Livnat Jerby, Alexandria Kluge, Abby Knecht, Monika Kowalczyk, John Kwon, Travis Law, Jane Lee, Joshua Levin, Robert Majovski, Kendra Mar, Christoph Muus, Sam Myers, Anna Neumann, Jackson Nyman, Yaara Oren, Tamara Ouspenskaia, Jenna Pfiffner, Dariusz Przybylski, Samantha Riesenfeld, Christopher Rodman, Noga Rogel, Orit Rosen, Geoffrey Schiebinger, Jonathan Schmid-Burgk, Kristine Schwenck, Karthik Shekhar, Meromit Singer, Marcin Tabaka, Dima Ter-Ovanesyan, Timothy Tickle, Itay Tirosh, Michael Tsabar, Josh Weinstein

