

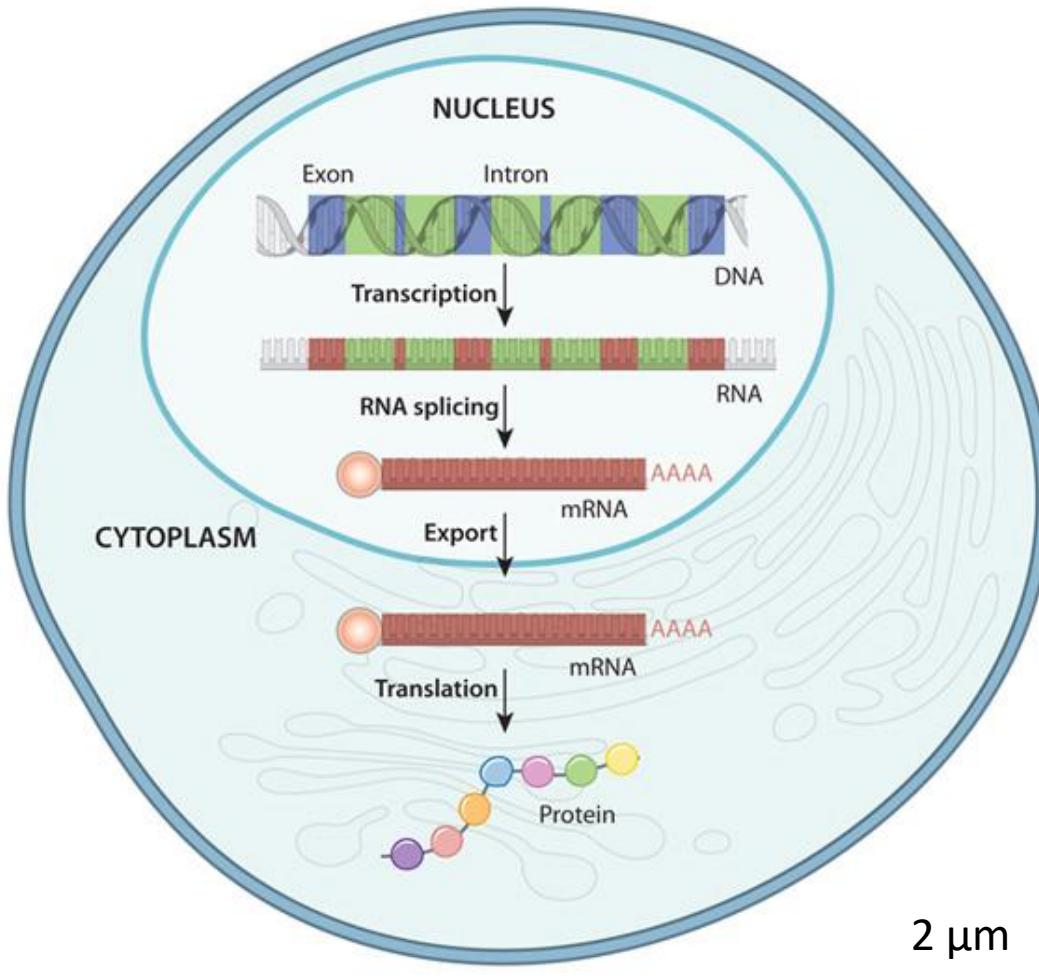
Lec 7. Single cell analysis

Zida Li
Associate Professor

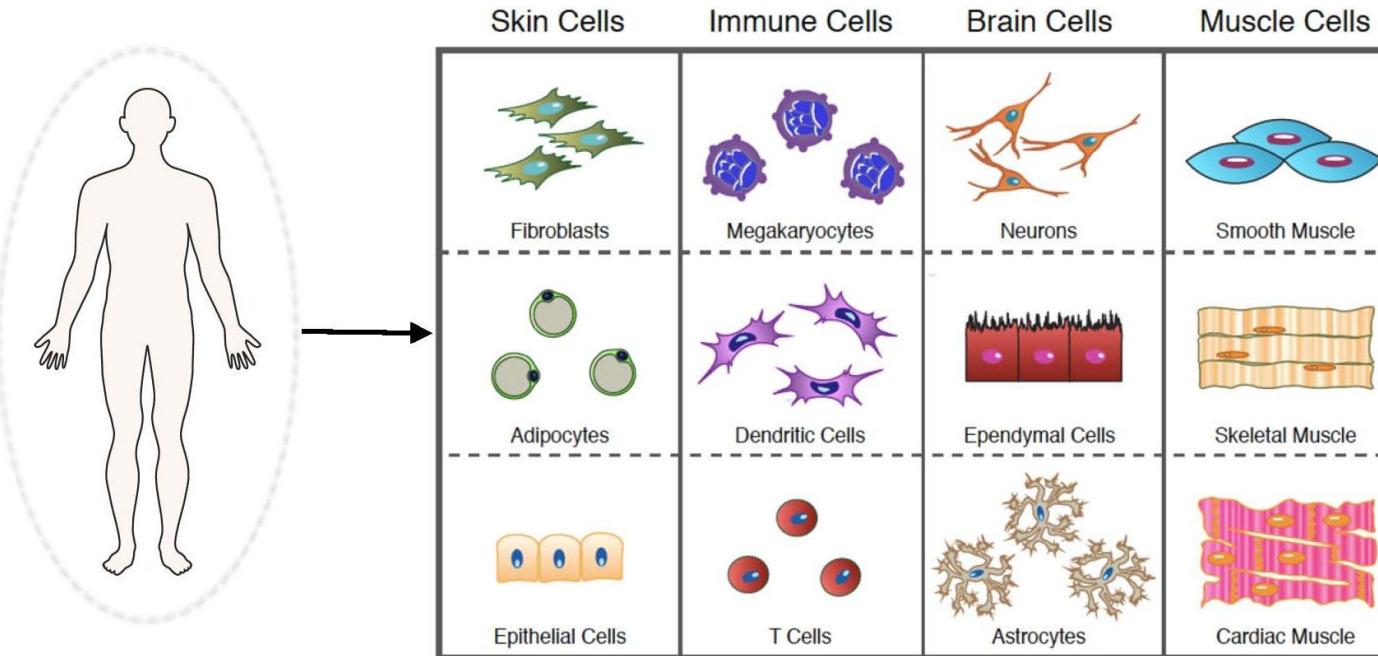


深圳大学
SHENZHEN UNIVERSITY

The basic components and central laws of cells



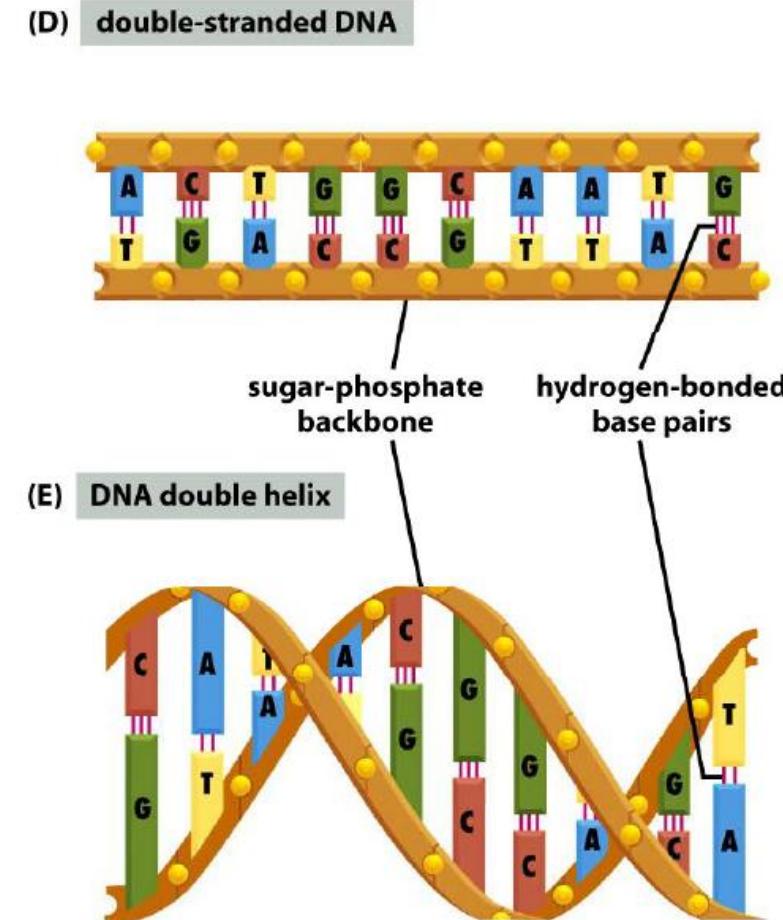
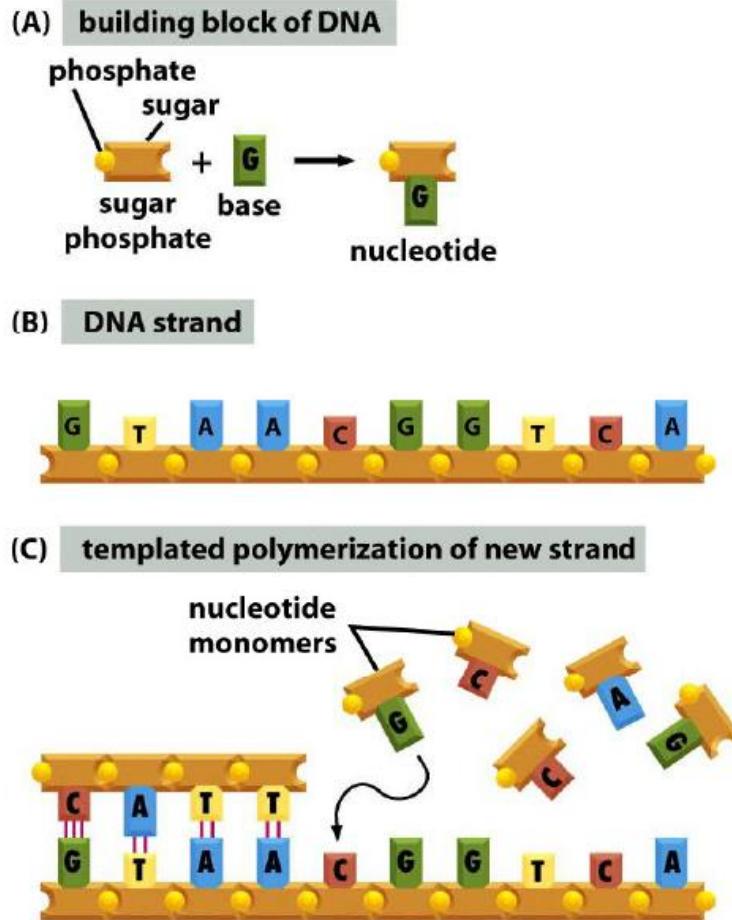
Cells are the core building blocks of the human body



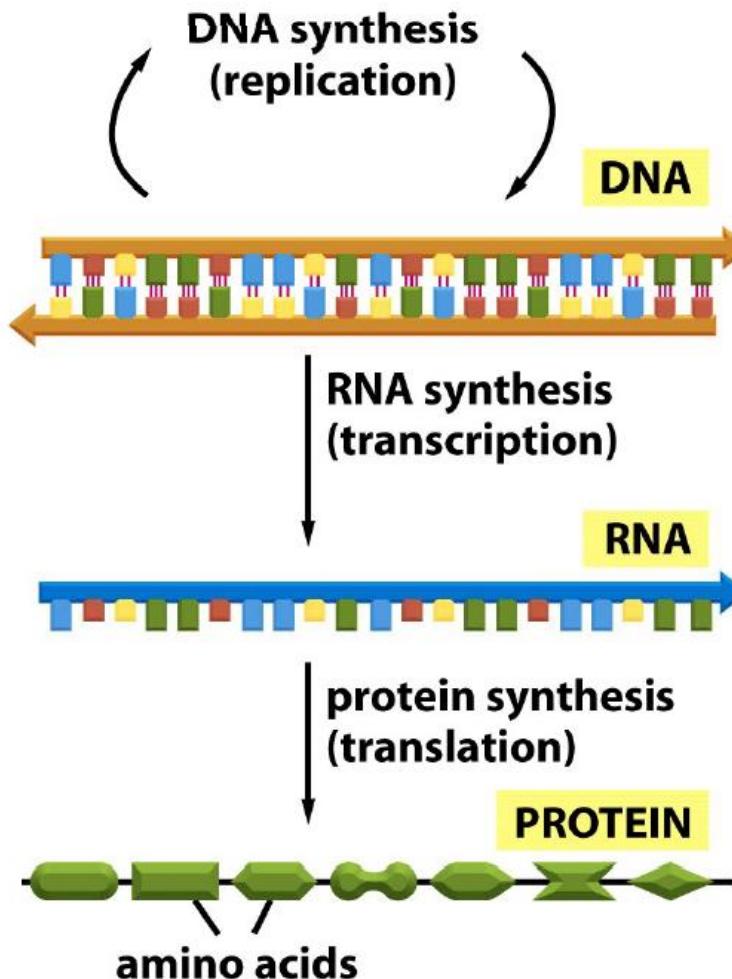
Cells are classified by characteristic molecules, structures, and functions

How did this happen?

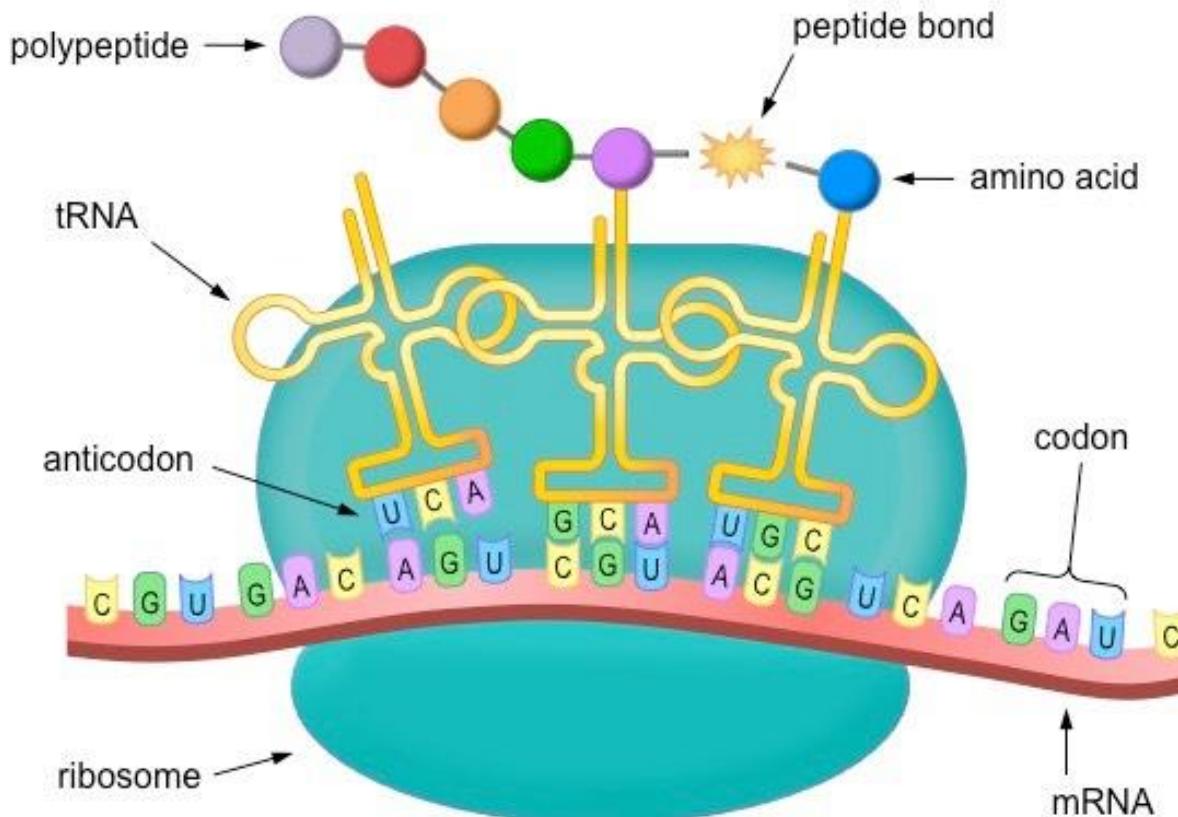
Molecular structure of DNA



Molecular mechanisms of transcription and translation



Molecular mechanisms of transcription and translation



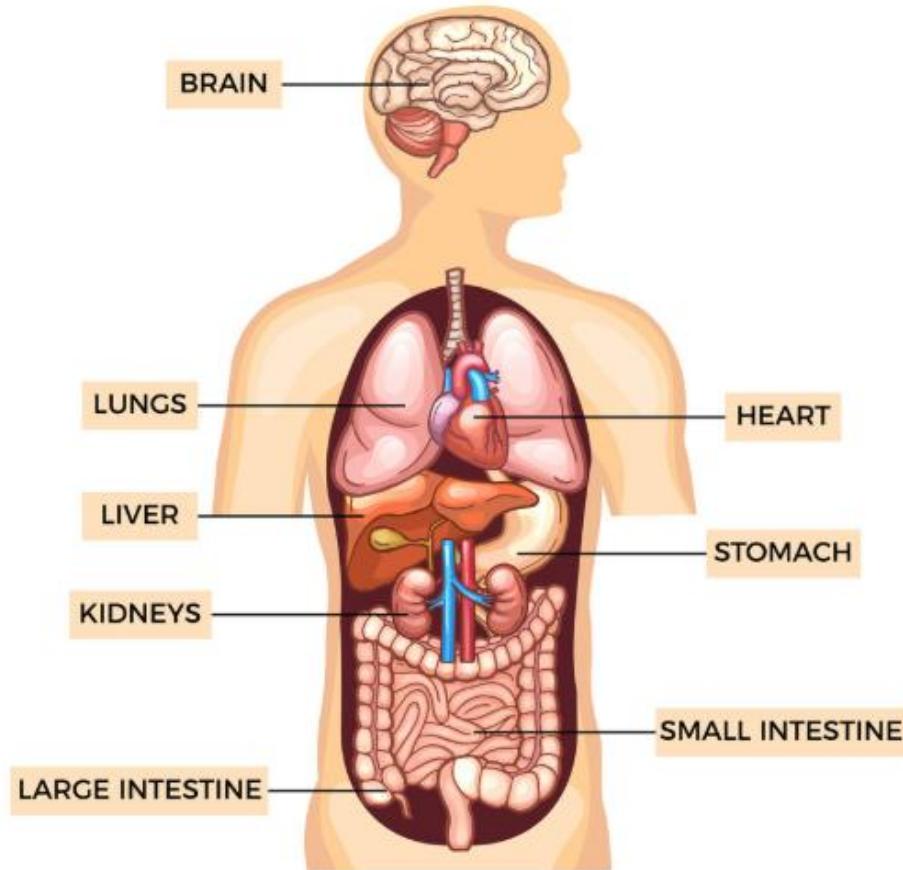
Organs made up of different cells make up the human body.

INTERNAL STRUCTURE OF THE HUMAN BODY

~30 trillion cells
30 万亿

Trillion
Billion
Million
Thousand

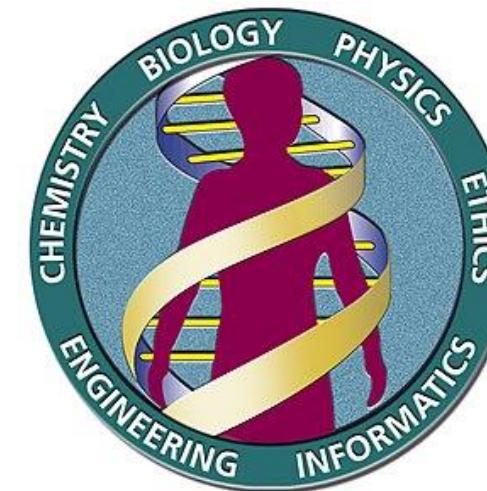
Milli-
Micro-
Nano-
Pico-
Femto-



人类基因组计划 (Human Genome Project)

- 1990-2003
- It remains the world's largest collaborative biological project.[\[1\]](#)
- the United States, the United Kingdom, Japan, France, Germany and China.
- Beijing Genomics Institute (华大基因研究院)/Human Genome Center, Institute of Genetics, Chinese Academy of Sciences, Beijing, China
- The estimated cost for generating that initial 'draft' human genome sequence is ~\$300 million worldwide, of which NIH provided roughly 50-60%.[\[2\]](#)

The screenshot shows the BBC News homepage. At the top, there is a navigation bar with links for Home, News (which is highlighted in red), Sport, Radio, TV, Weather, and Languages. Below this is a red banner with the BBC News logo and a link to 'One-Minute World News'. The main content area has a red background. At the top of this area, it says 'Last Updated: Monday, 14 April 2003 16:01 GMT 17:01 UK'. Below this, there are links for 'E-mail this to a friend' and 'Printable version'. The main headline reads 'Human genome finally complete' in bold black text. Below the headline, it says 'By Ivan Noble BBC News Online science staff'. A sub-headline states: 'The biological code crackers sequencing the human genome have said they have finished the job - two years ahead of schedule.' On the left side of the main content area, there is a sidebar with links for 'News Front Page', 'Africa', 'Americas', 'Asia-Pacific', 'Europe', and 'Middle East'.



Logo of the Human Genome Project

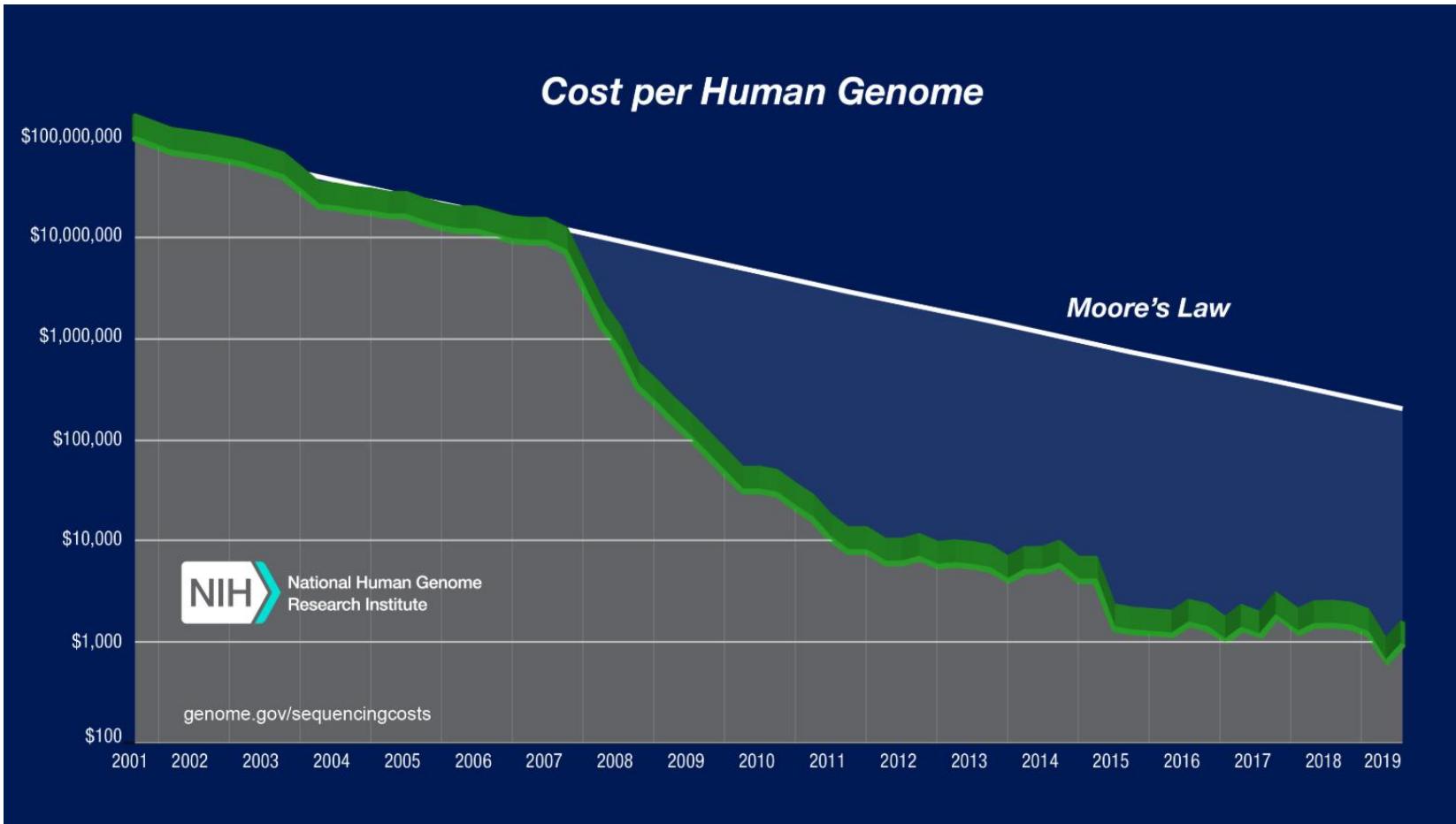
Key findings in the Human Genome Project

- There are approximately 22,300^[36] protein-coding genes in human beings, the same range as in other mammals.
- The human genome has significantly more segmental duplications (nearly identical, repeated sections of DNA) than had been previously suspected.^{[37][38][39]}
- At the time when the draft sequence was published, fewer than 7% of protein families appeared to be vertebrate specific.^[40]

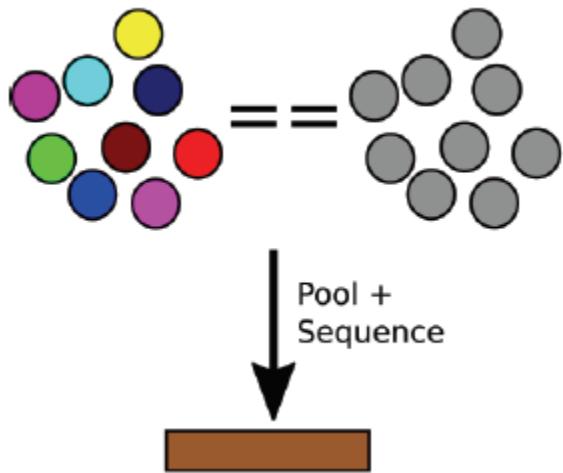
Source:



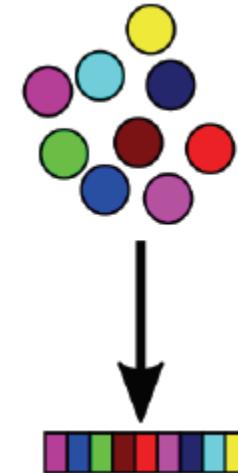
测序成本的降低



单细胞测序可以揭示细胞间的异质性

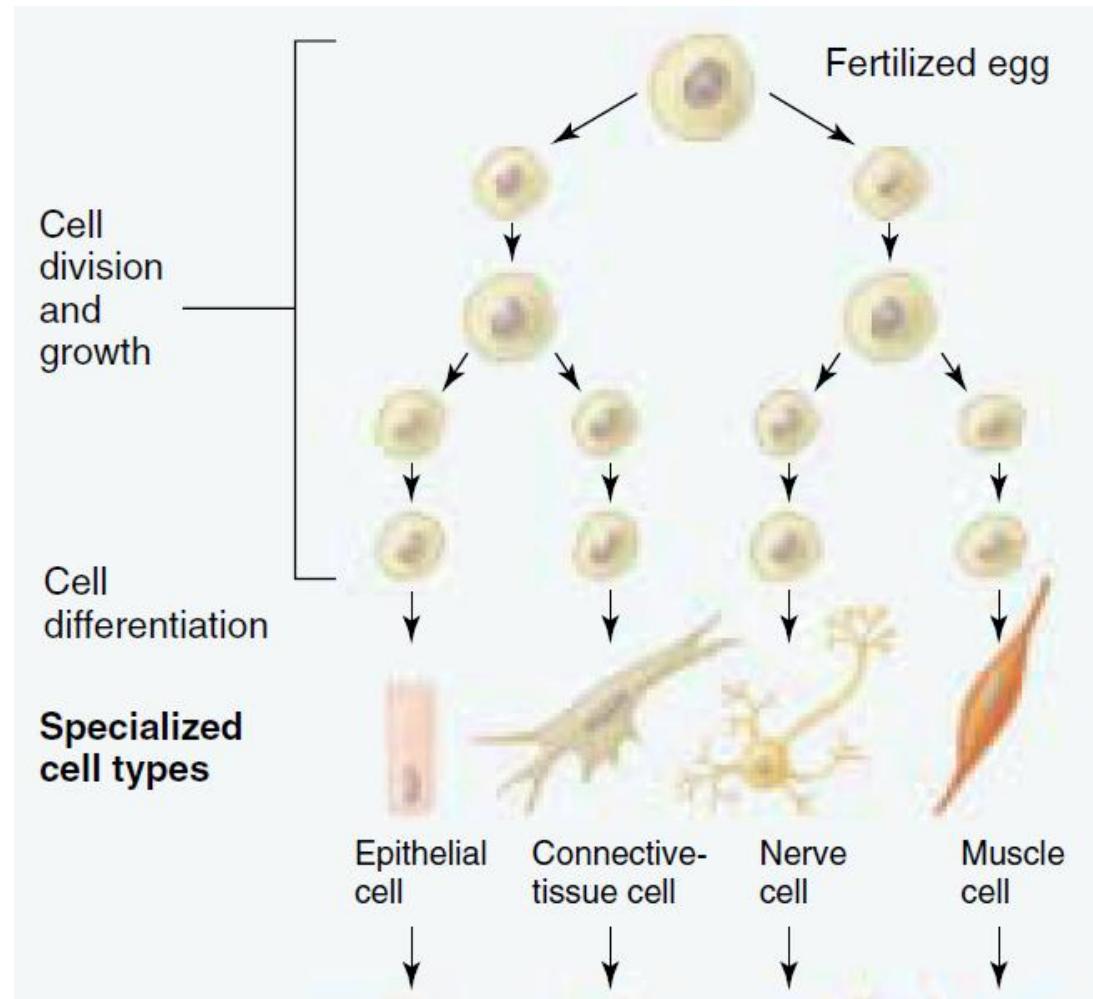


Bulk measurement provides ensemble average.

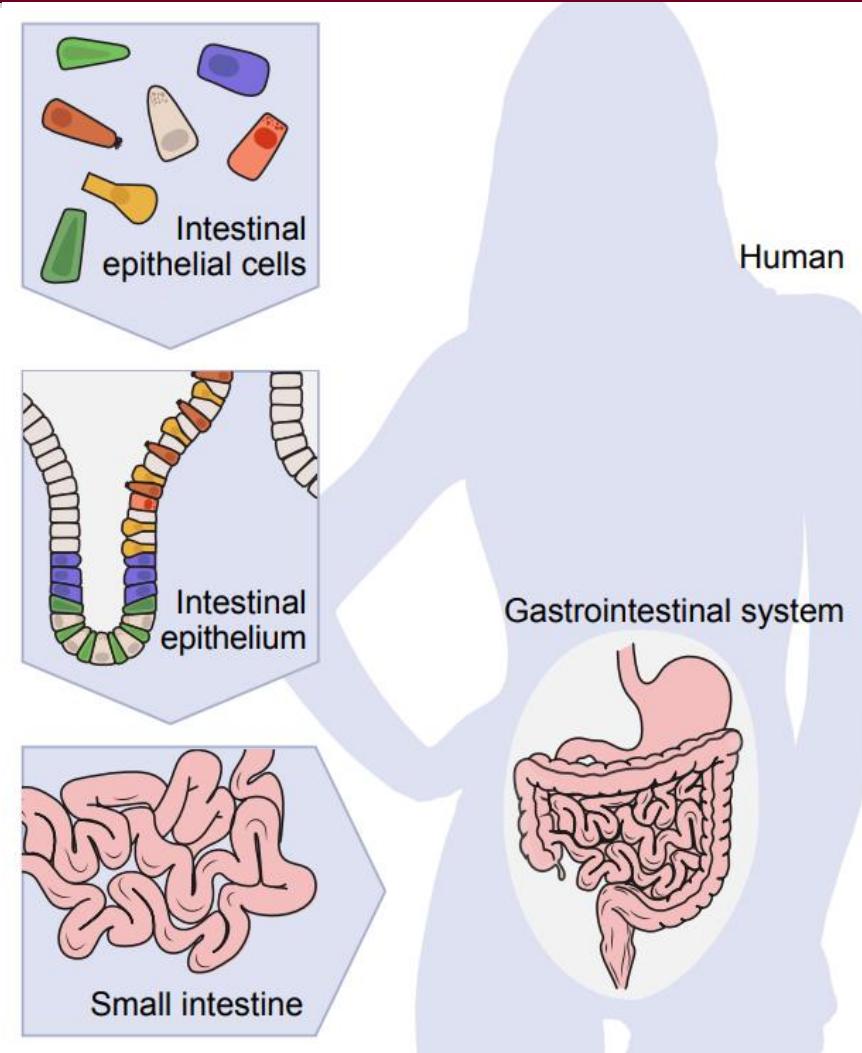


Single-cell measurement gauges individual cells.

器官是由不同的组织构成



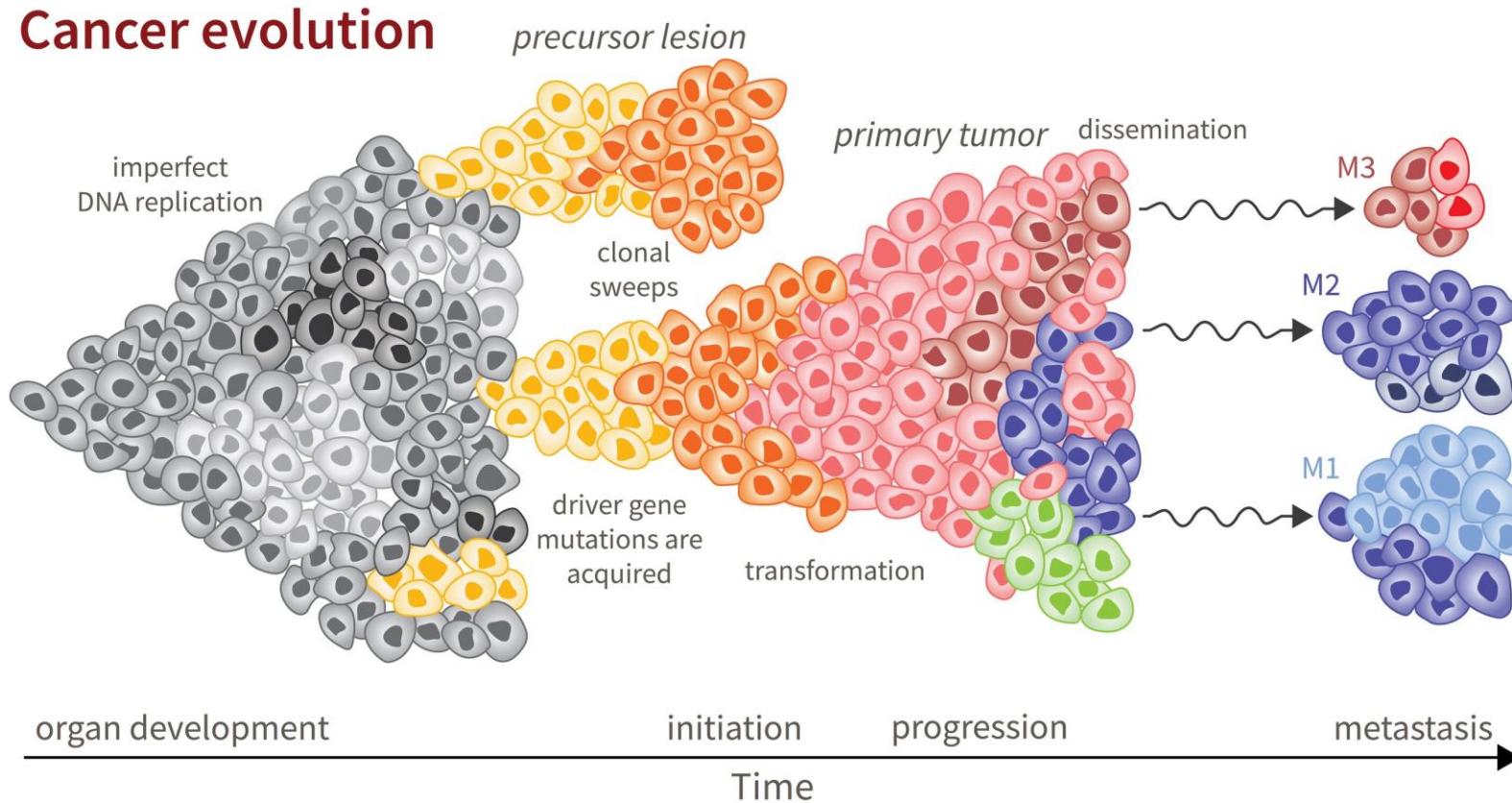
单细胞测序可以揭示细胞间的异质性



单细胞测序的应用

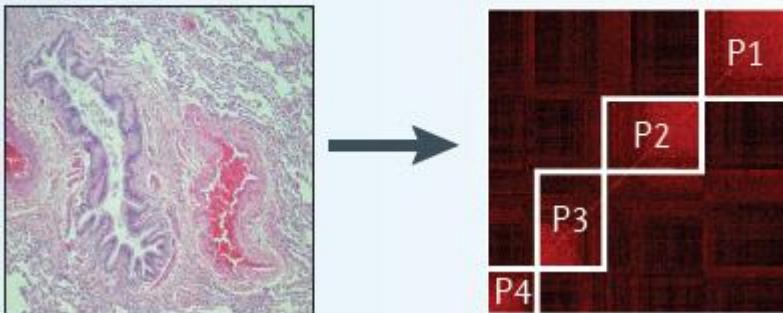
- Cancer
- Developmental biology
- Immunology
- Neurobiology
- Plant biology

Cancer:

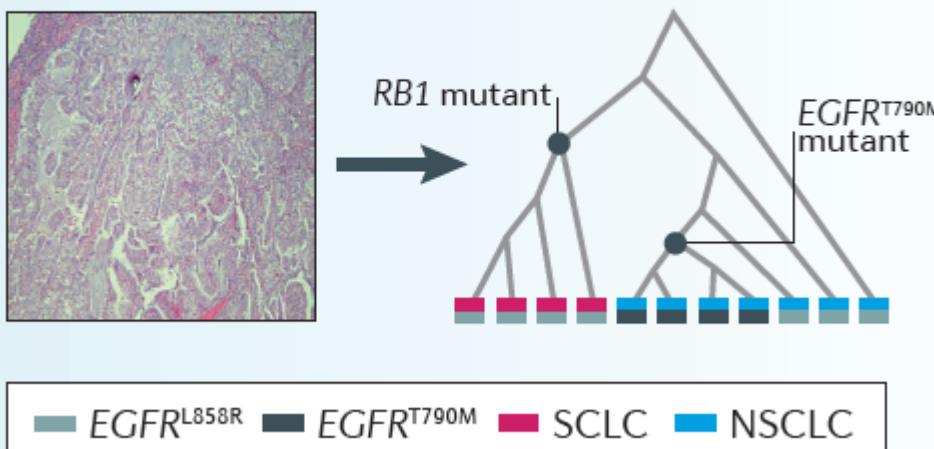


单细胞测序在癌症中的应用

- a** Single-cell RNA-seq to identify lung tissue cell types and their associated biomarkers



- b** Single-cell genome sequencing followed by lineage tracing to identify the genetics of tumour evolution



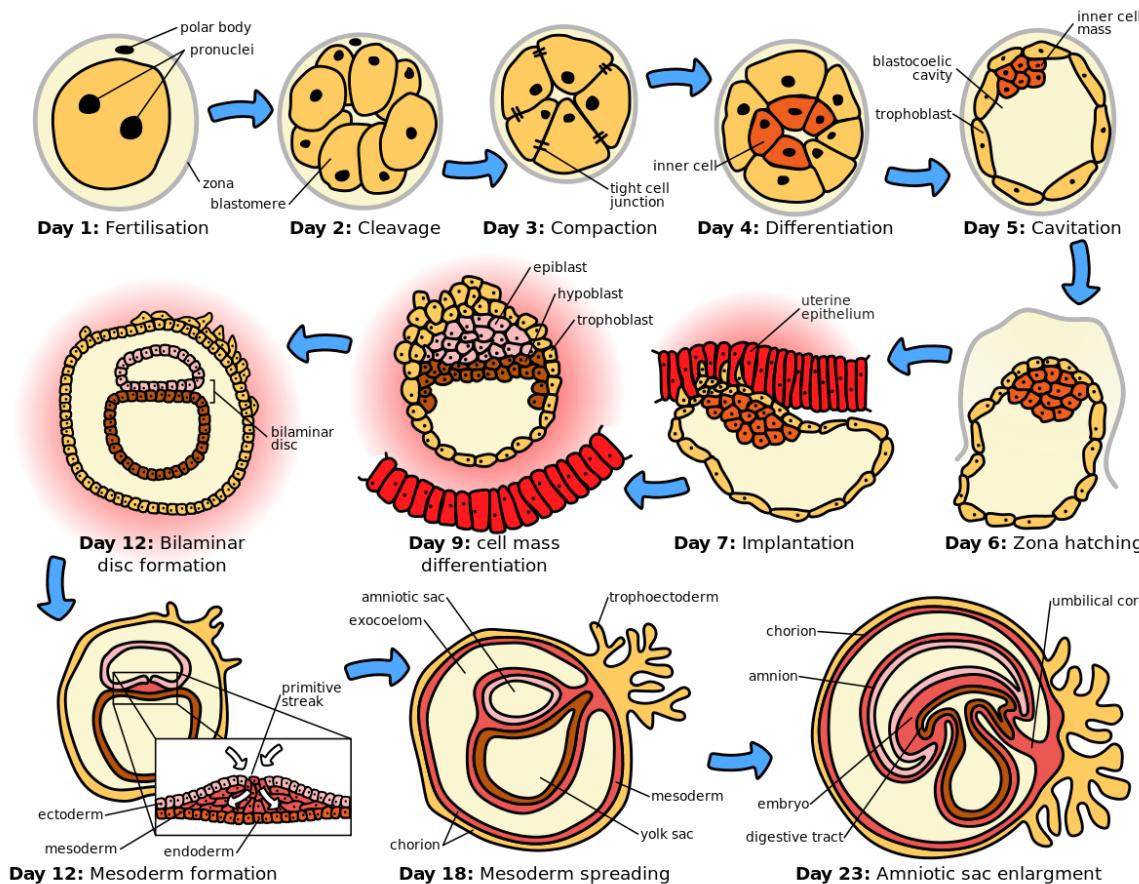
OPINION

Unravelling biology and shifting paradigms in cancer with single-cell sequencing

Timour Baslan and James Hicks

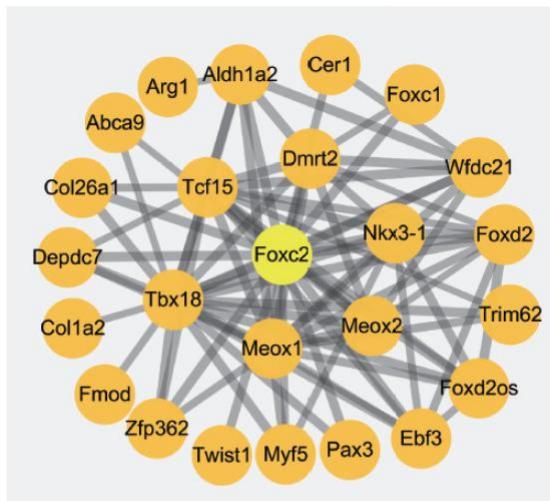
Developmental Biology

- Lineage tracing in early embryonic development

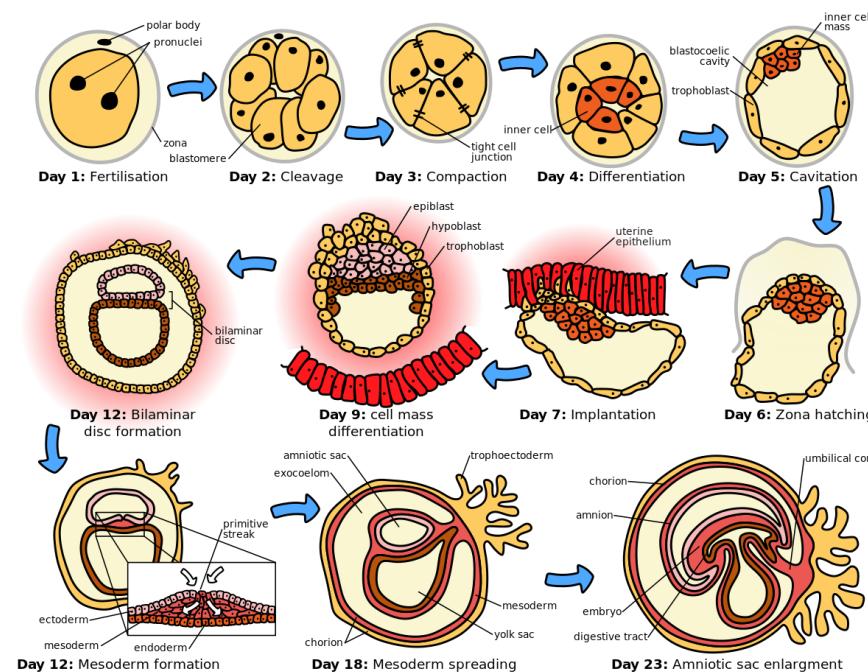


Developmental Biology

- Lineage tracing in early embryonic development
- Gene regulation of developmental processes
- Regenerative medicine

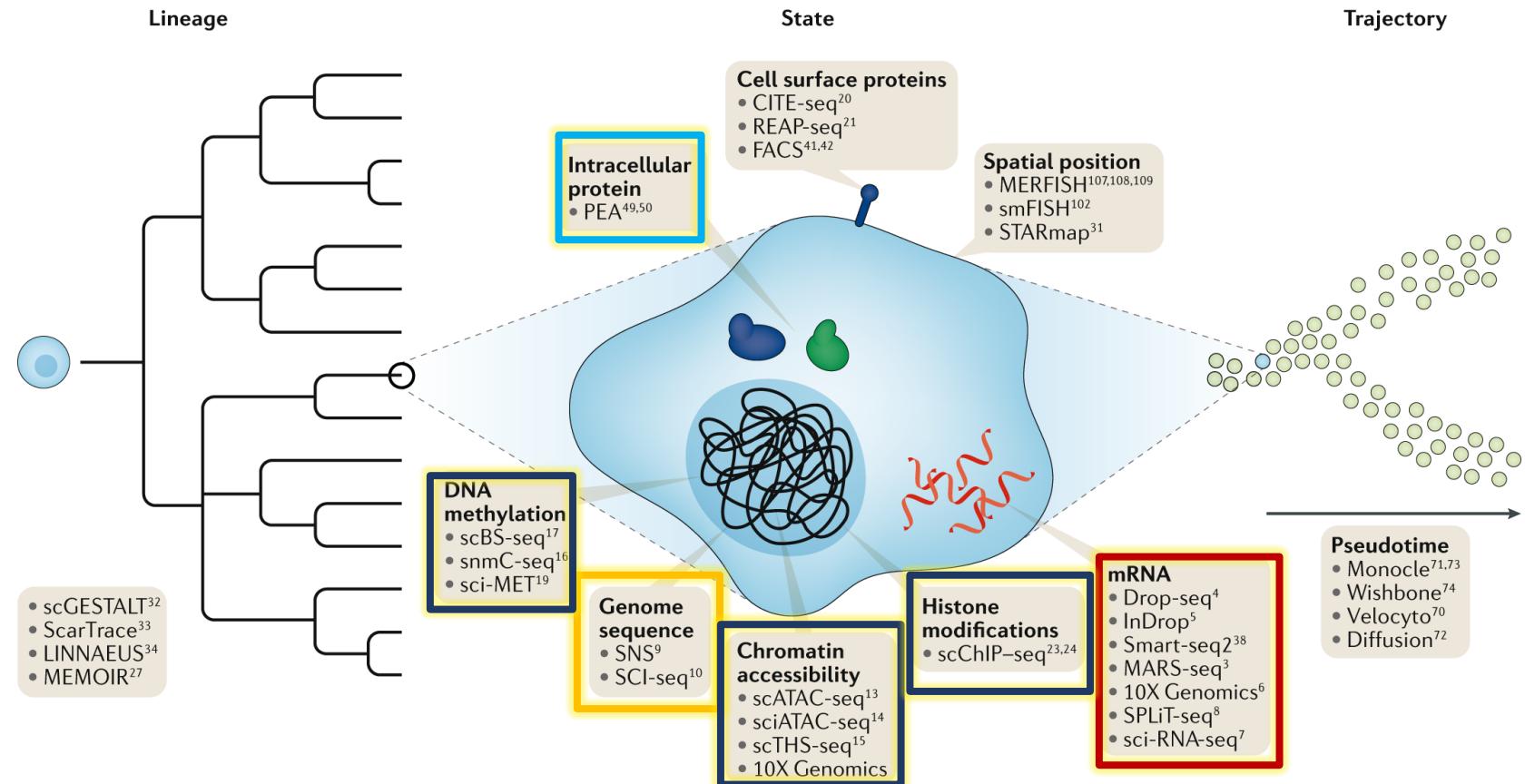


Gene regulation network
(SCENIC, WGCNA, etc.)

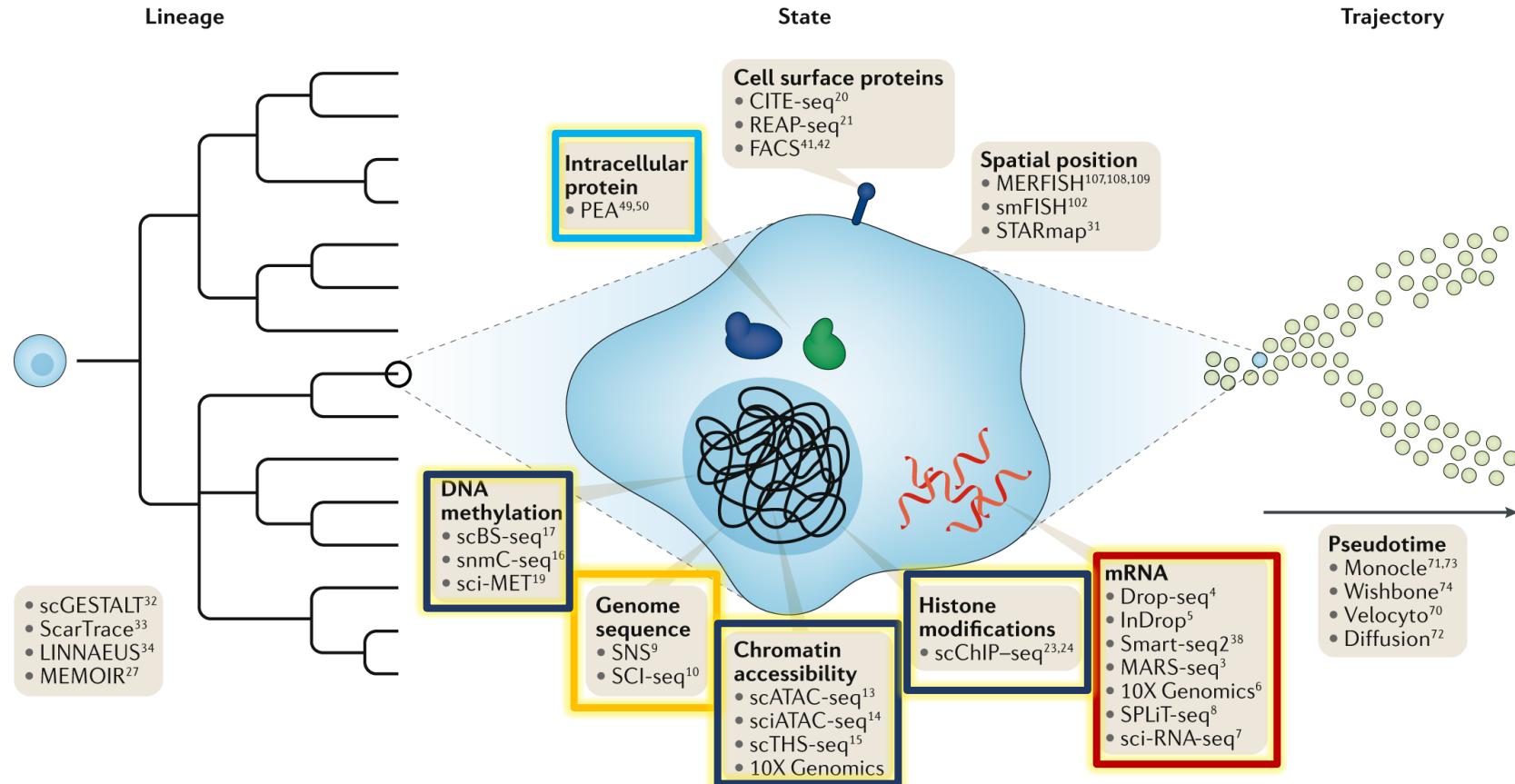


Single cell analysis includes the analysis of genome, epigenome, transcriptome, metabolome, and proteome.

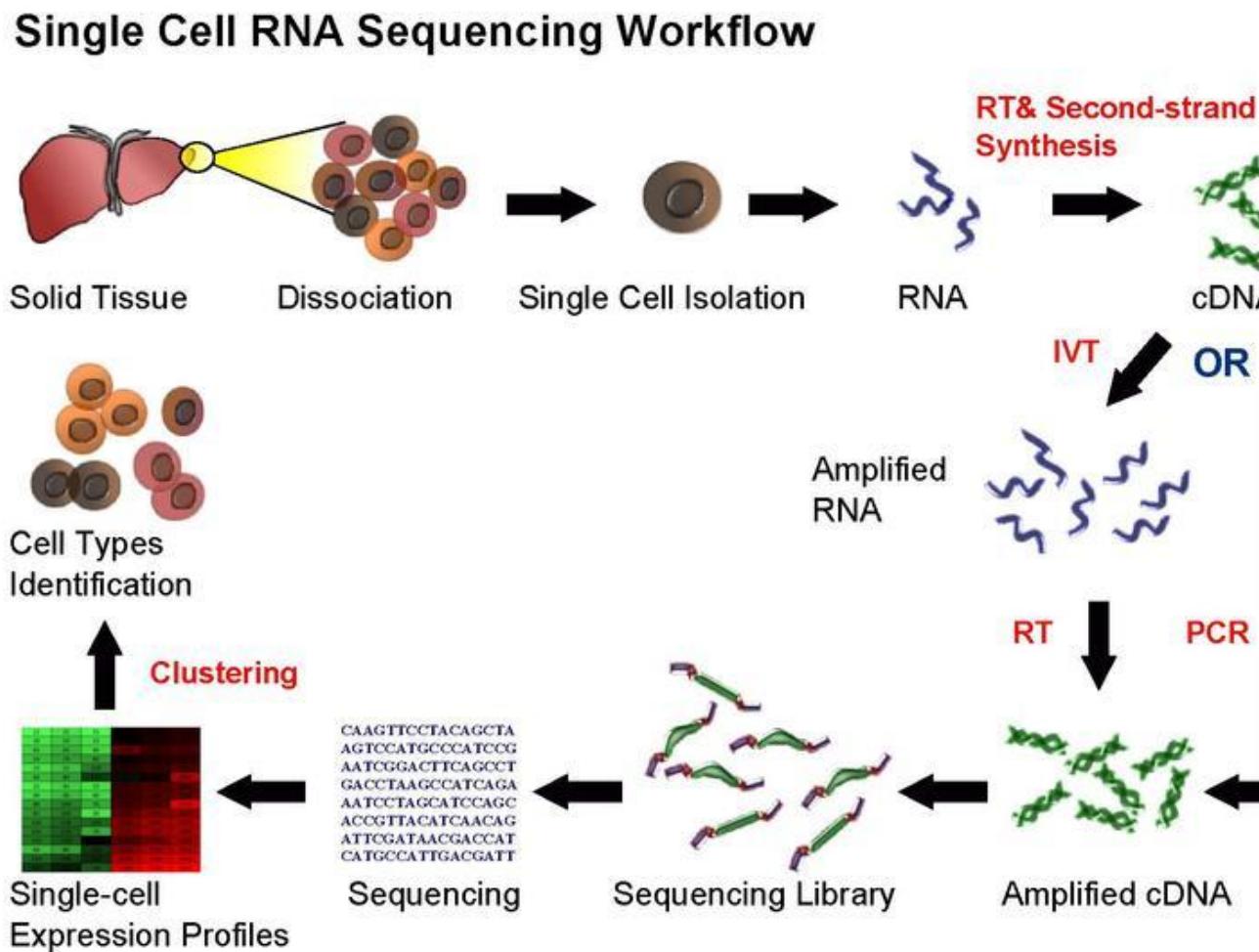
- 转录组 (transcriptome)
- 基因组 (genome)
- 表观基因组 (epigenome)
 - 甲基化
 - Histone
- 空间组学 (spatial-omics)



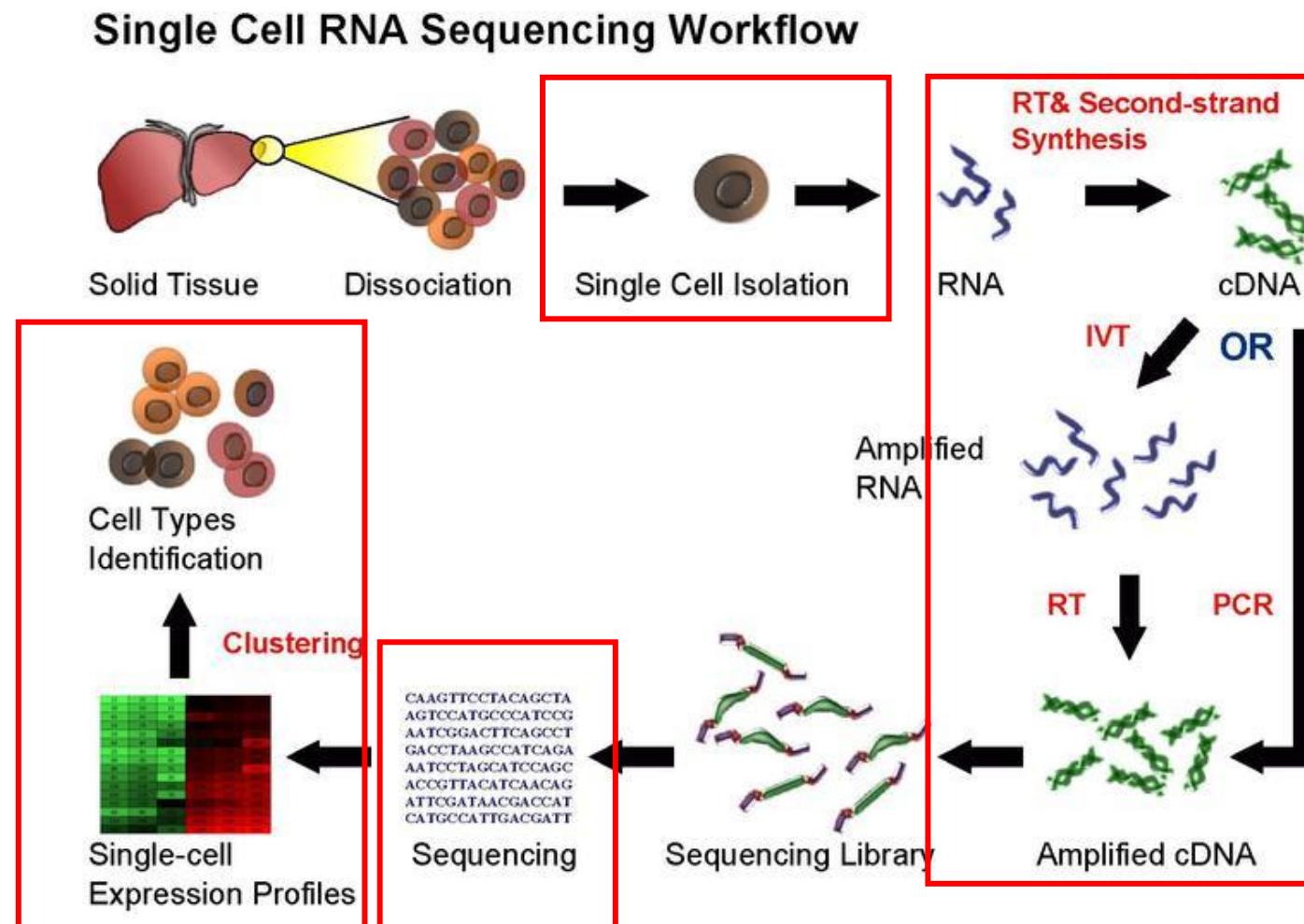
Single cell analysis includes the analysis of genome, epigenome, transcriptome, metabolome, and proteome.



The basic process of single-cell transcription sequencing



The basic process of single-cell transcription sequencing



scRNA-seq

Chemistry:

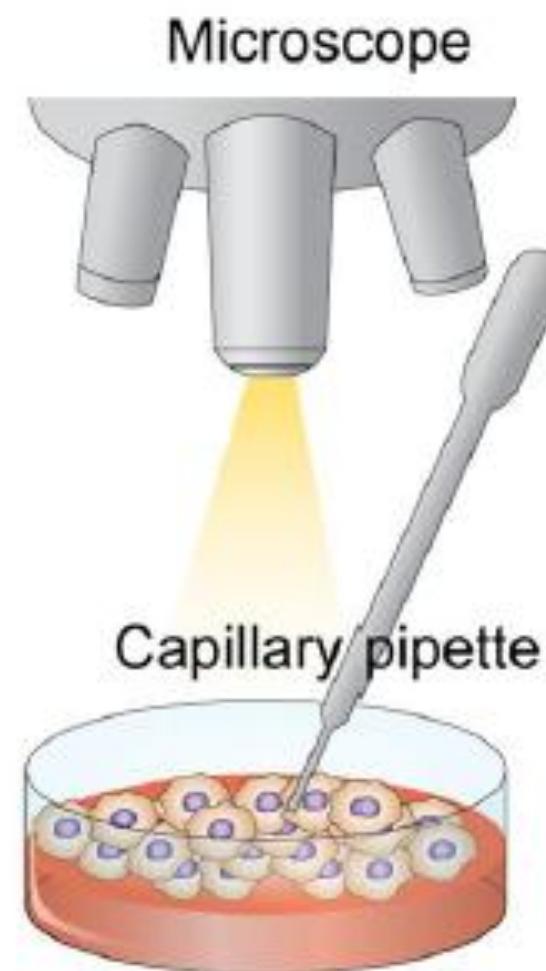
- 2009: Tang et al. (Nature Methods)
- 2012: Smart-seq
 - 2013: Smart-seq2

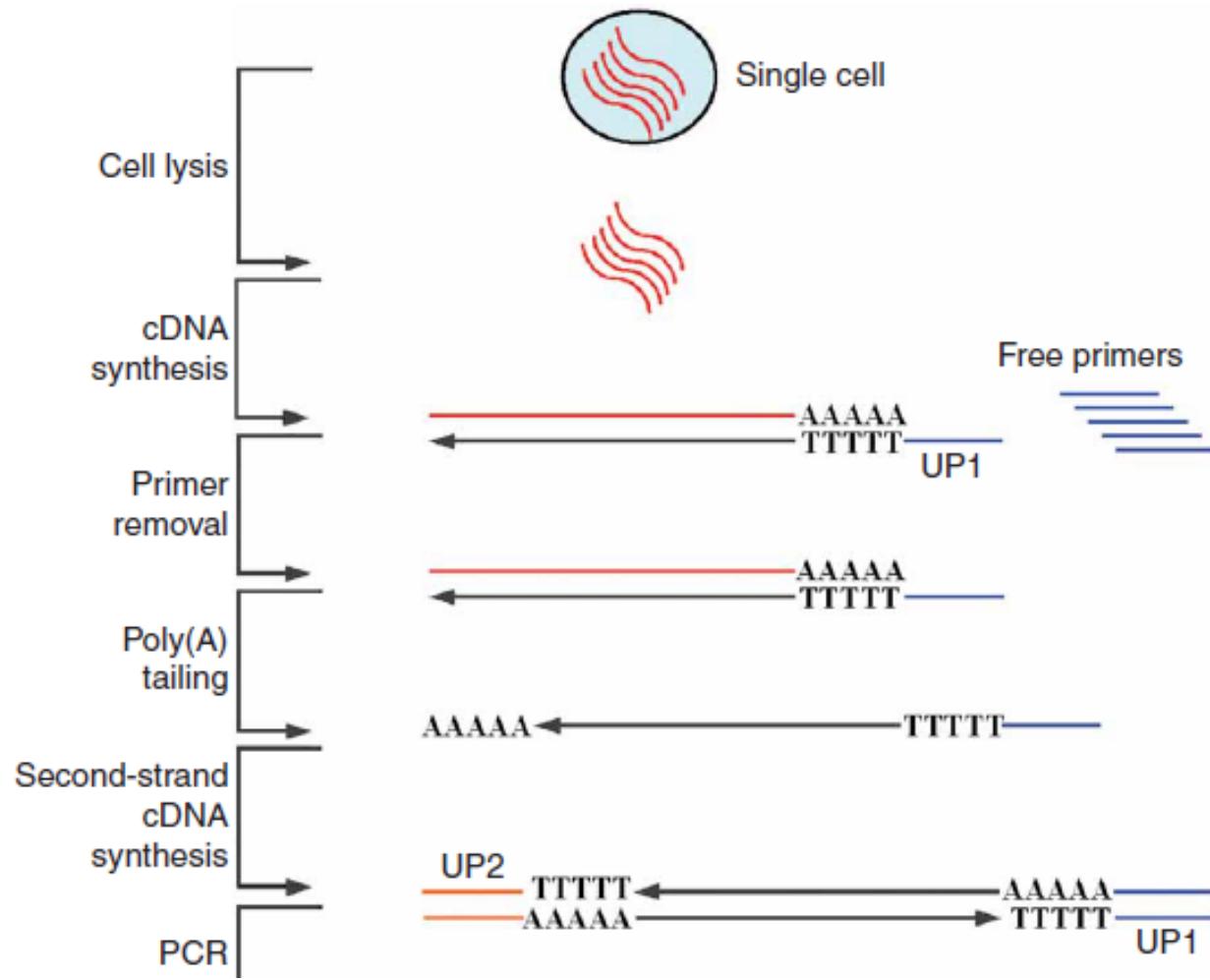
Automation:

- 2014: Fluidigm C1
- 2015: Drop-seq and inDrop, 10X Genomics

The first successful single-cell transcriptome sequencing

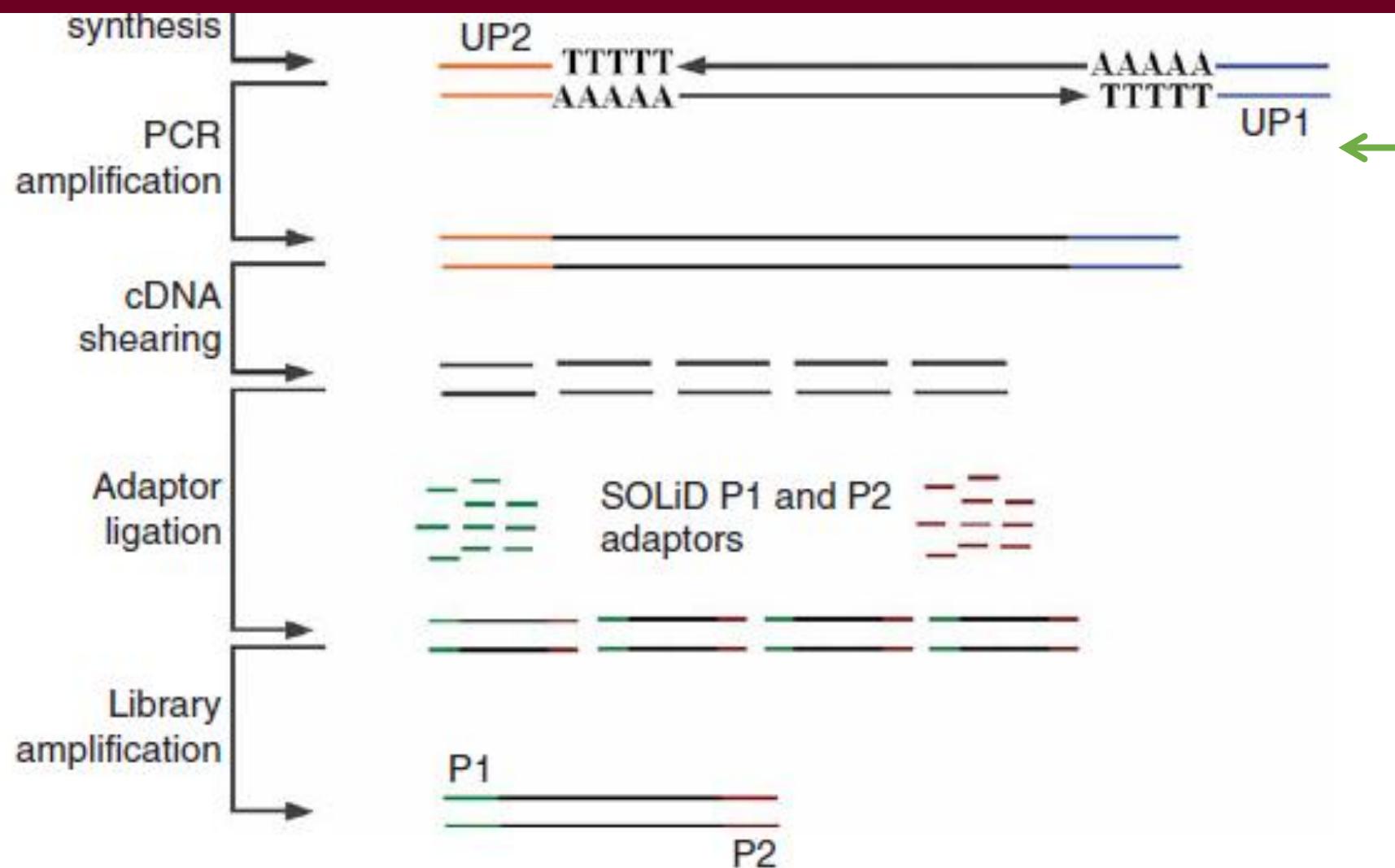
- Nature Methods, 2009
- Background:
- NGS generally requires micrograms of RNA counts.
- This requirement is difficult to meet in some important situations, such as:
- Early differentiation of mouse embryos (~30 cells)
- Embryonic stem cell culture: many subclasses
- It is necessary to study single-cell transcription sequencing





Degradation of unreacted primers by Exonuclease I





Challenge (pre-2009)

Needed micrograms of RNA for library prep

High loss from multi-step protocols

Poor coverage due to inefficient RT

No universal primers for amplification

Incompatibility with NGS library prep

Tang et al. (2009) Solution

Whole-transcriptome PCR amplification from picograms

One-tube enzymatic workflow minimizing loss

Anchored oligo(dT) priming for complete mRNA capture

UP1/UP2 anchor system enabling exponential amplification

Amine-blocked primers + SOLiD-compatible workflow

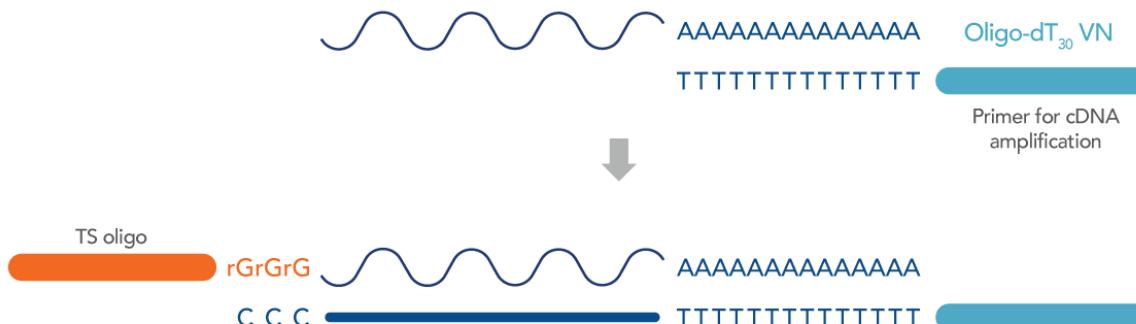
Smart-seq

- **SMART-seq**: Switching mechanism at the 5' end of the RNA transcript
- Rickard Sandberg, Karolinska Institute, Sweden
- That's what **SMART** stands for:
 - Switching Mechanism At the RNA Template.
- This made the method:
 - More efficient,
 - More reproducible,
 - And less biased than the poly(A)-tailing + second-strand synthesis used by Tang et al.

Cell-derived poly(A)-tailed mRNA

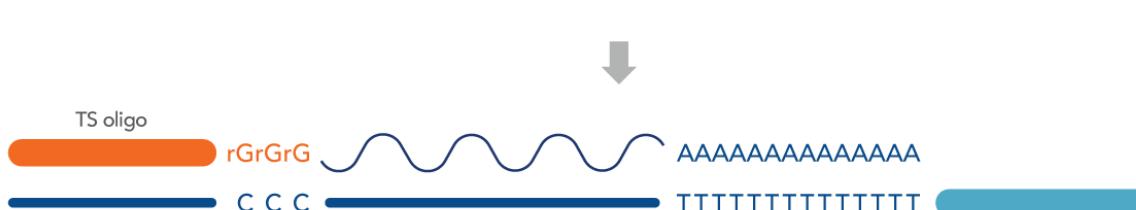
Reverse transcription

(1st strand synthesis)



Template switching

(from mRNA to TS oligo)



cDNA amplification



- **Step 1. Reverse transcription with an oligo(dT) primer**
- Each mRNA's **poly(A) tail** is primed by an **oligo(dT)** primer containing:
 - A **universal adapter sequence** at its 5' end (for later PCR),
 - A stretch of **T residues** at its 3' end (for annealing to the mRNA's poly(A) tail).
- This ensures that only mRNAs (not rRNA or tRNA) are reverse transcribed.

- **Step 2. Template switching by reverse transcriptase**
- The reverse transcriptase (typically **Moloney Murine Leukemia Virus (MMLV)**, a special “SMART” version) copies the mRNA into cDNA until it reaches the 5' end.
- Upon reaching the end, the enzyme **adds a few non-templated cytosines** to the 3' end of the newly made cDNA.
- A **synthetic template-switching oligo (TSO)**, containing:
 - A few **guanosines (GGG)** at its 3' end,
 - And a known sequence at its 5' end, base-pairs with these cytosines, allowing the reverse transcriptase to **“switch templates”** and extend the cDNA, adding the TSO sequence to its 5' end.

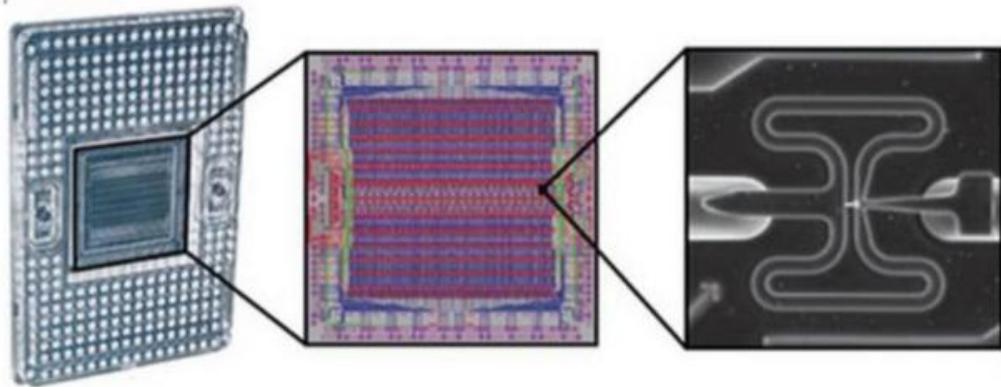
- **Step 3. PCR amplification**
- Using primers that recognize those universal sequences, the cDNA pool is **uniformly amplified**.
- Since the entire transcript (from 5' cap to 3' poly(A) tail) is now represented, you get **full-length coverage**, not just 3' ends.

What SMART-seq improved over Tang 2009

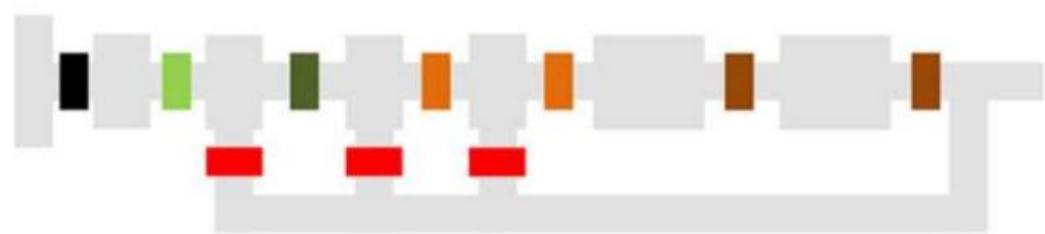
Feature	Tang et al. (2009)	SMART-seq (2012)
Second-strand synthesis	Poly(A) tailing + anchored oligo(dT) primer (UP2)	Template switching using reverse transcriptase
Adapters added	By tailing and primer design	During reverse transcription (built-in)
cDNA coverage	Partial (3' enriched, incomplete 5' ends)	Full-length , covering all exons and splice junctions
Steps and losses	Multi-step, enzyme-intensive	Single-tube, fewer losses
Quantitative bias	Moderate PCR bias	Lower bias, more linear amplification
Readout	SOLiD, short reads	Illumina, longer paired-end reads

Fluidigm C1

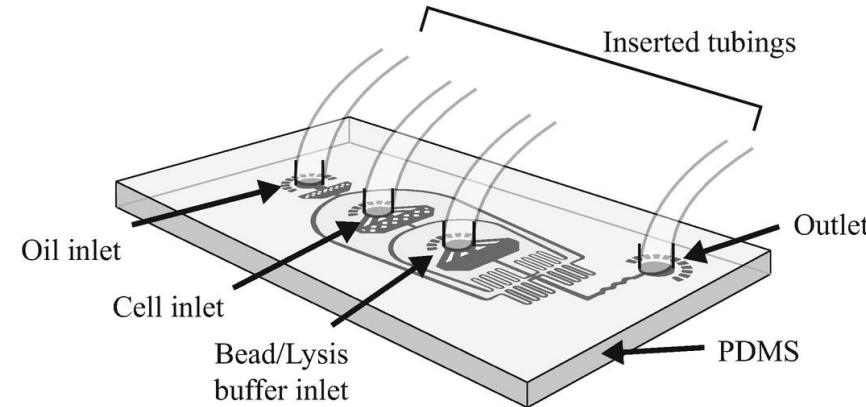
- The **Fluidigm C1 Single-Cell Auto Prep System** is a **microfluidic instrument** that isolates individual cells into tiny reaction chambers and automatically performs all the steps of single-cell cDNA preparation — from cell capture to amplification — with minimal user intervention.
- In essence, it's an *automated laboratory on a chip* designed for single-cell genomics.
- Published alongside the **SMART-seq** chemistry (and often used together), the C1 enabled the first **high-throughput, full-length scRNA-seq workflows**.

A**B**

Cell Capture	Cell Lysis	Cell Neutralization	Multiple Displacement Amplification (MDA)
4.5 nL	9 nL	9 nL + 9 nL	135 nL + 135 nL

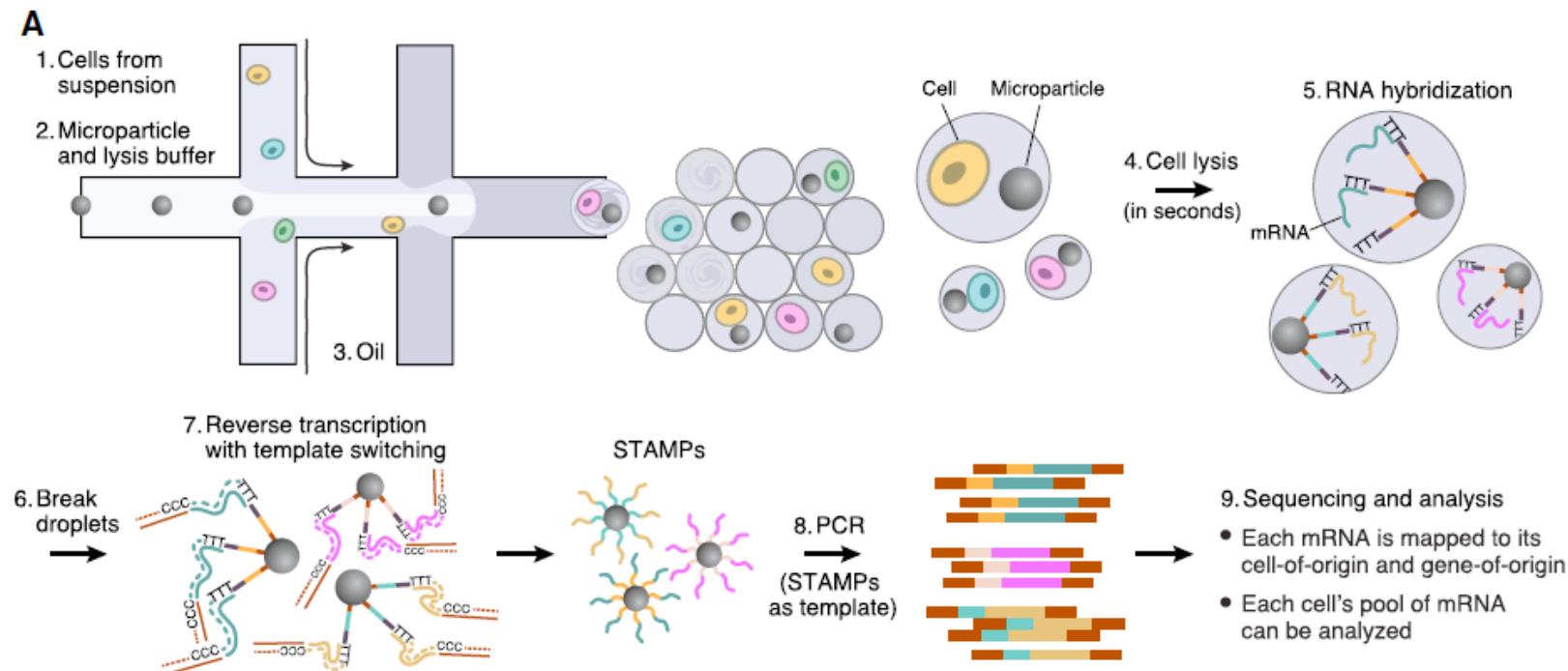


Drop-seq

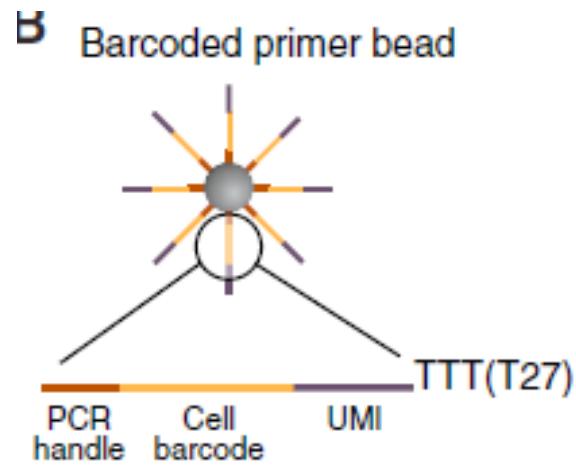


- Harvard Medical School
- Broad Institute of Harvard and MIT

Drop-seq

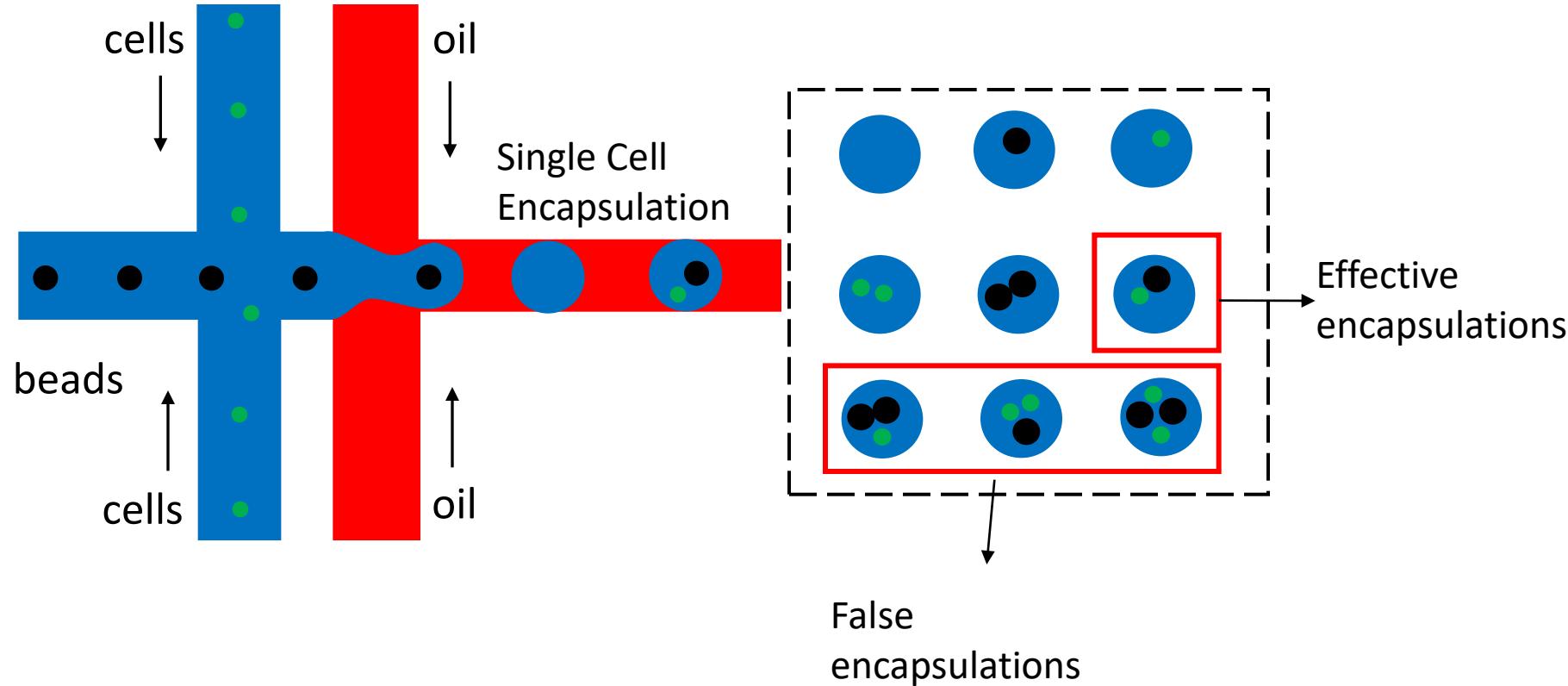


- Cell barcode
- UMI: Unique molecular identifier



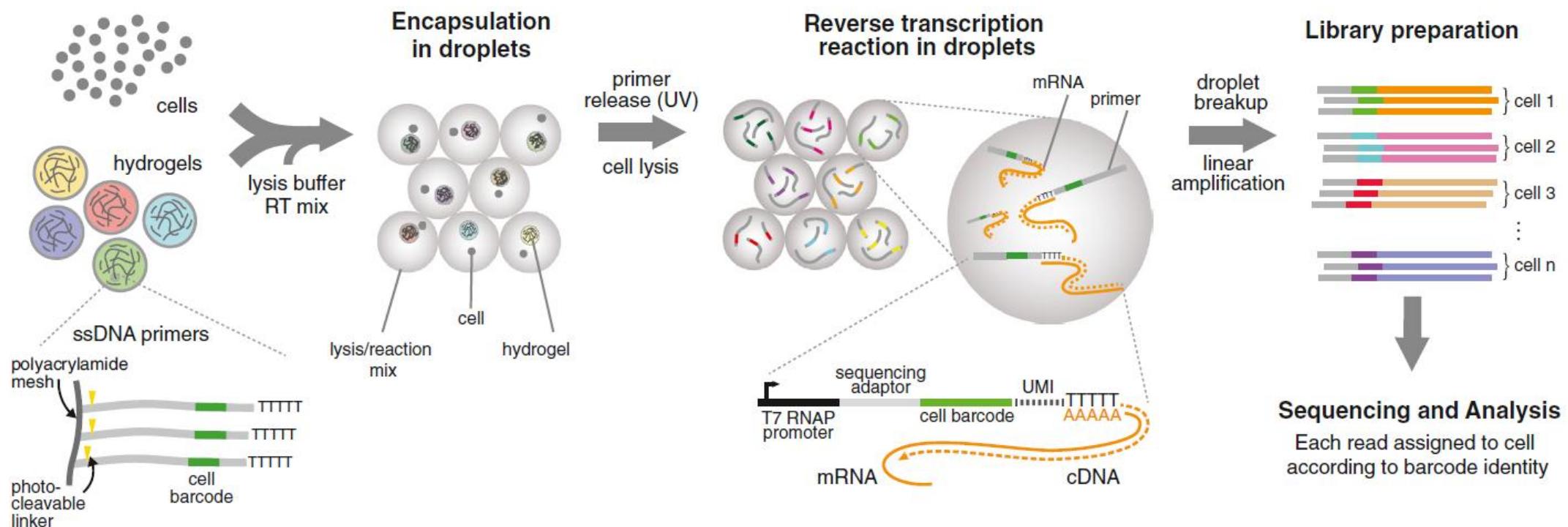
- 液滴微流控
 - ~500 液滴/s
- 分析了49k个细胞

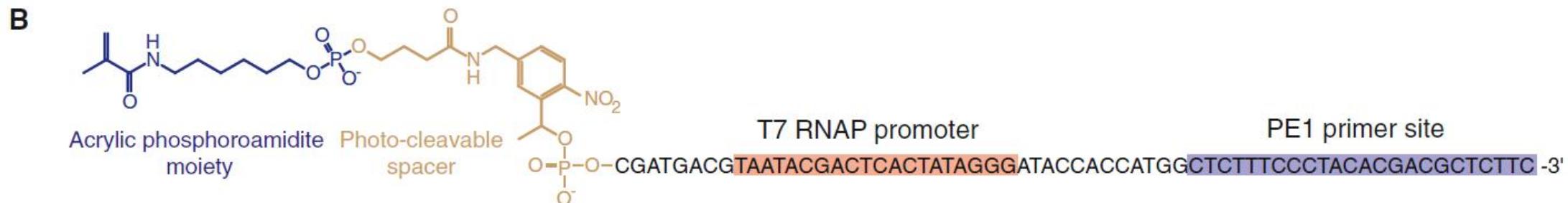
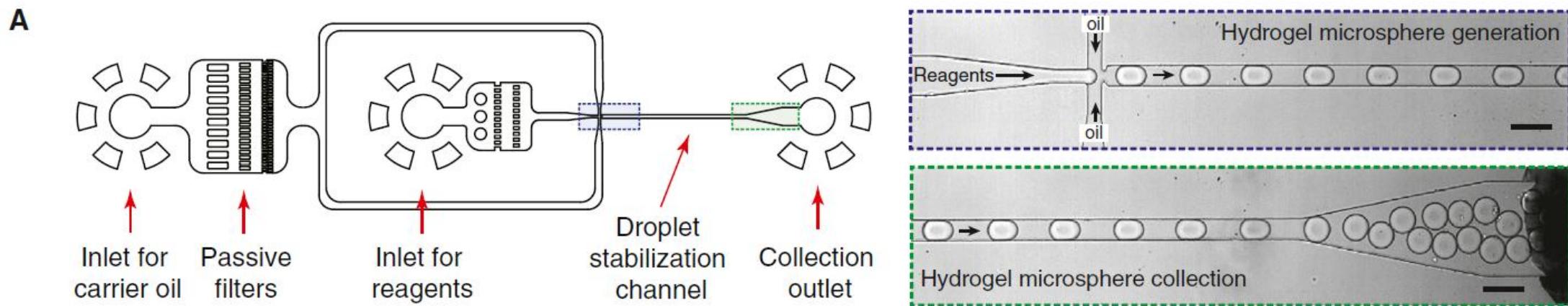
Limitations of Drop-seq

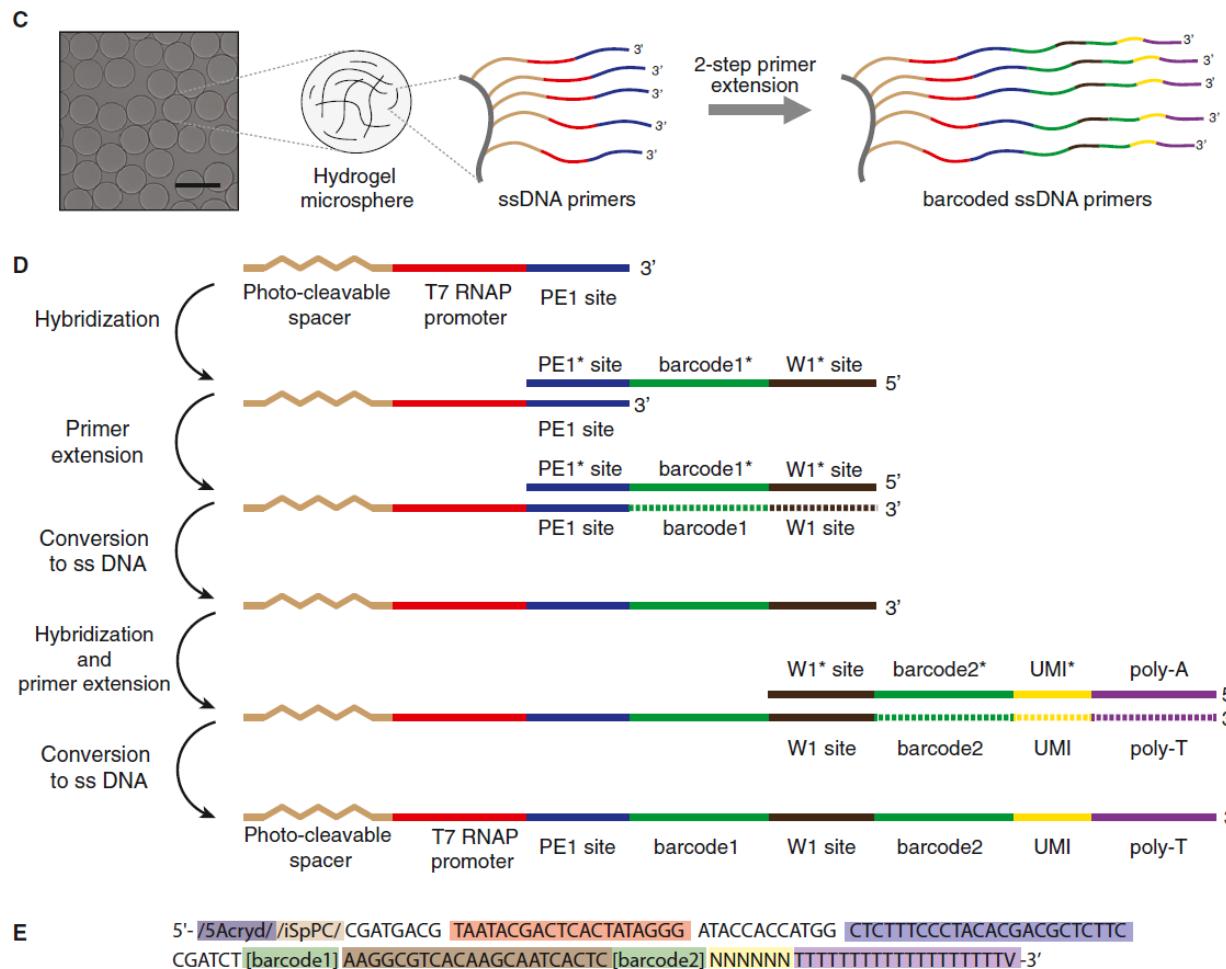


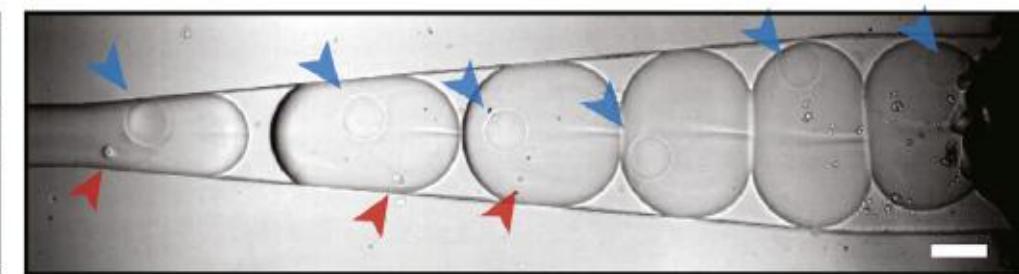
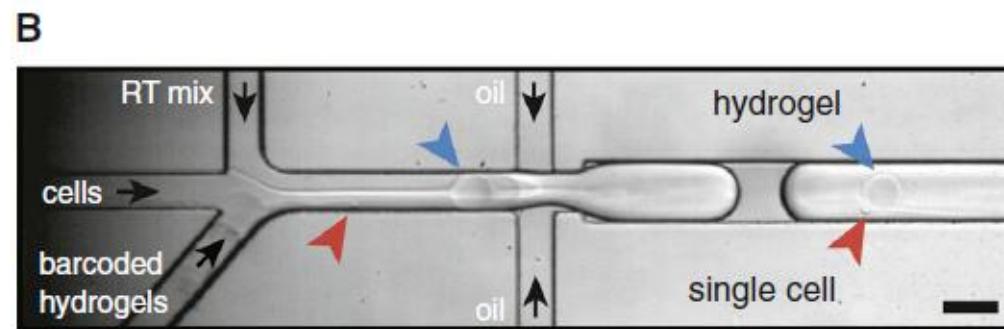
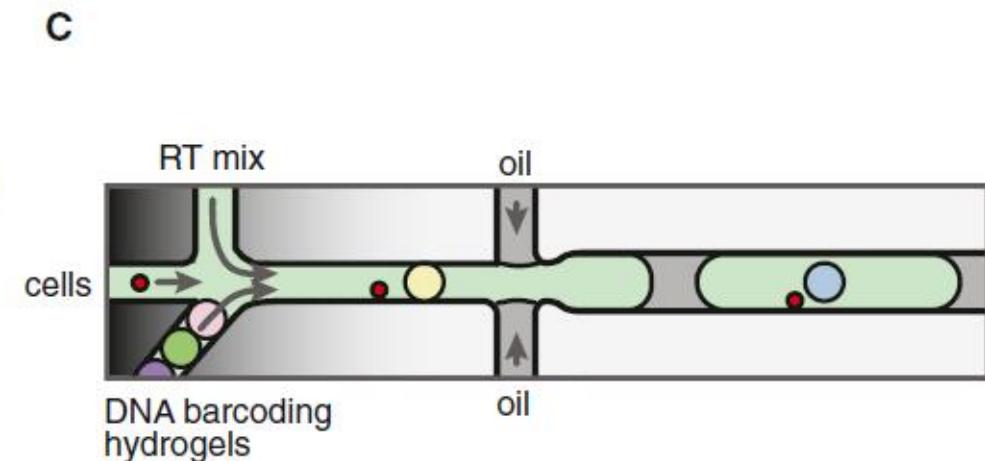
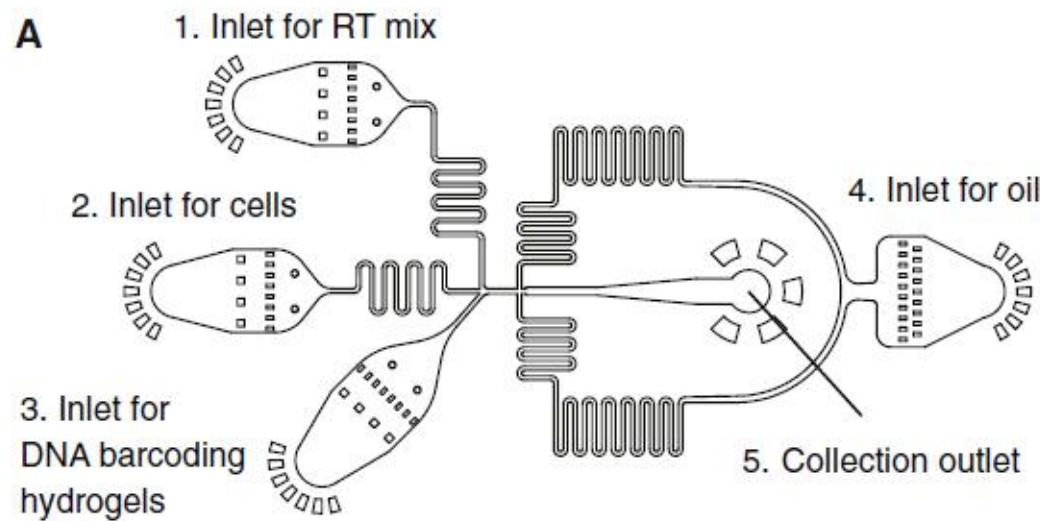
- Drawbacks: Low cell capturing efficiency.

inDrop

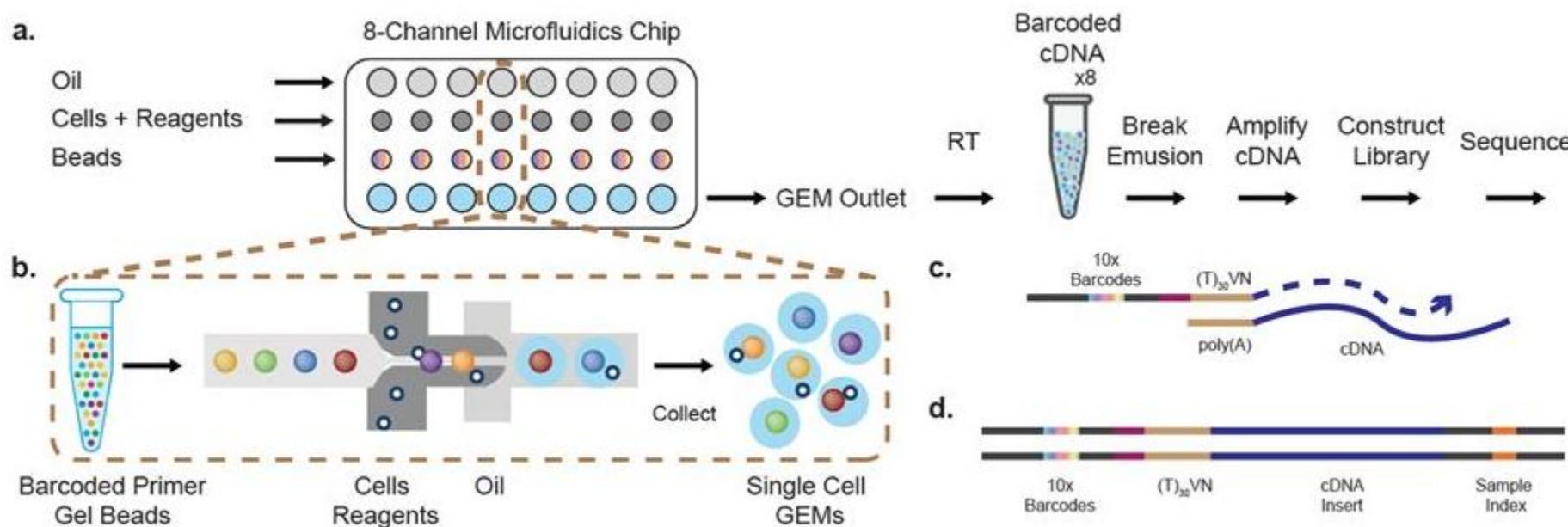


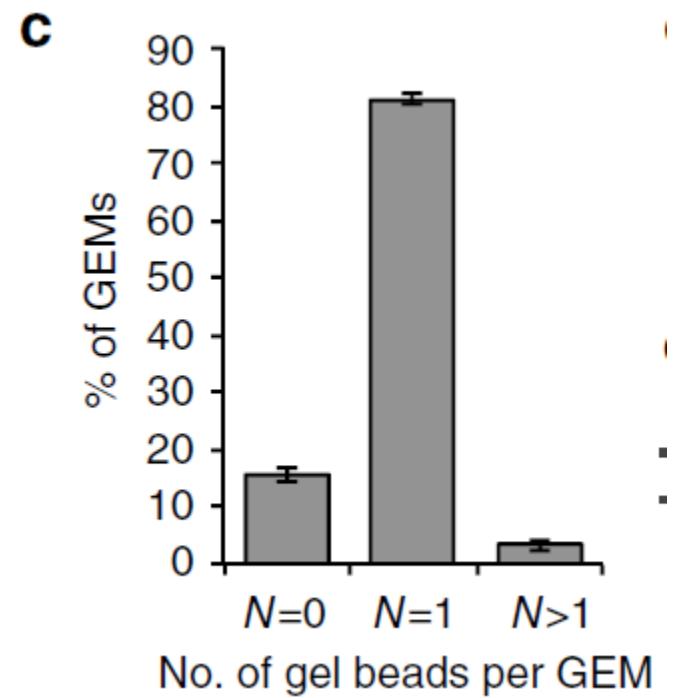
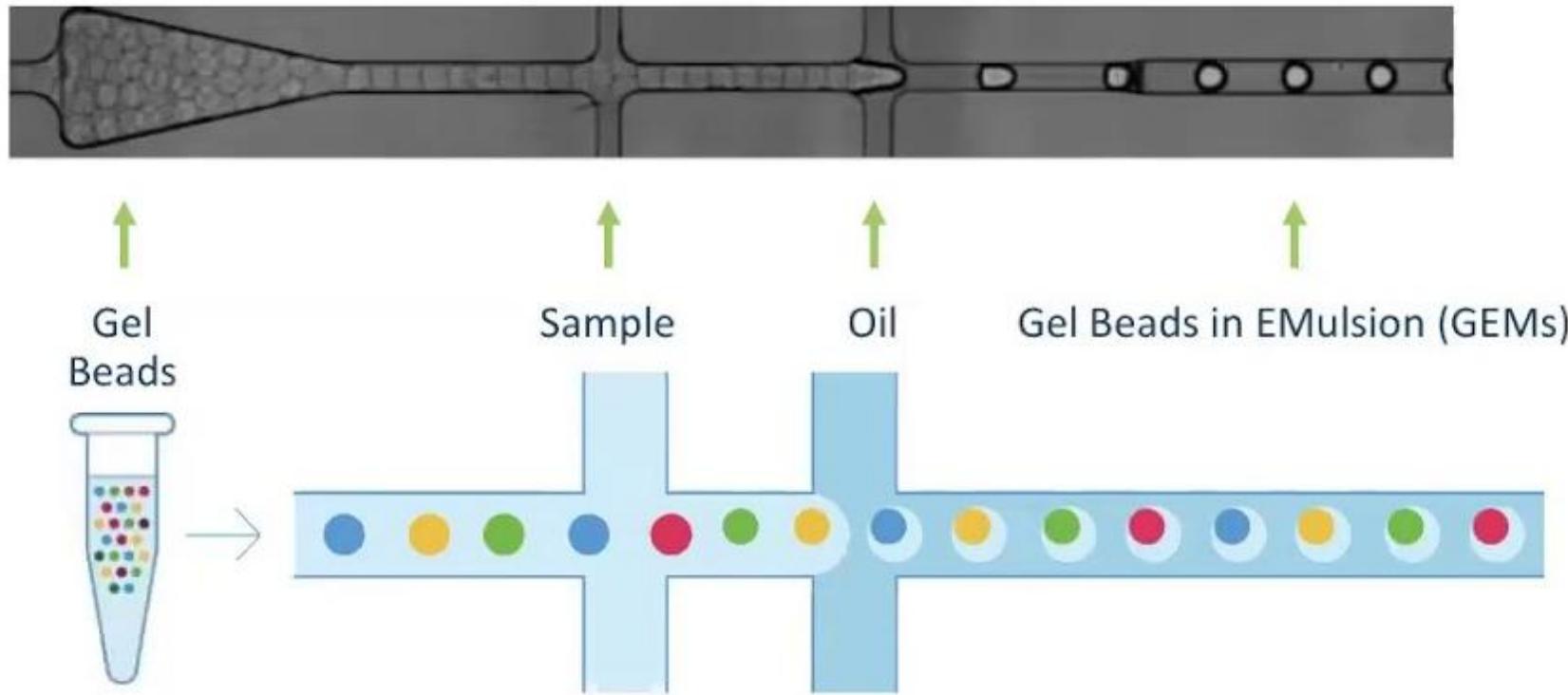






10X Genomics





Technology variations

- Spin-Drop
- ImgPico
- PIP-seq
- CP-seq
- CITE-seq
- SEC-seq



spinDrop: a droplet microfluidic platform to maximise single-cell sequencing information content

Received: 5 January 2023

Accepted: 21 July 2023

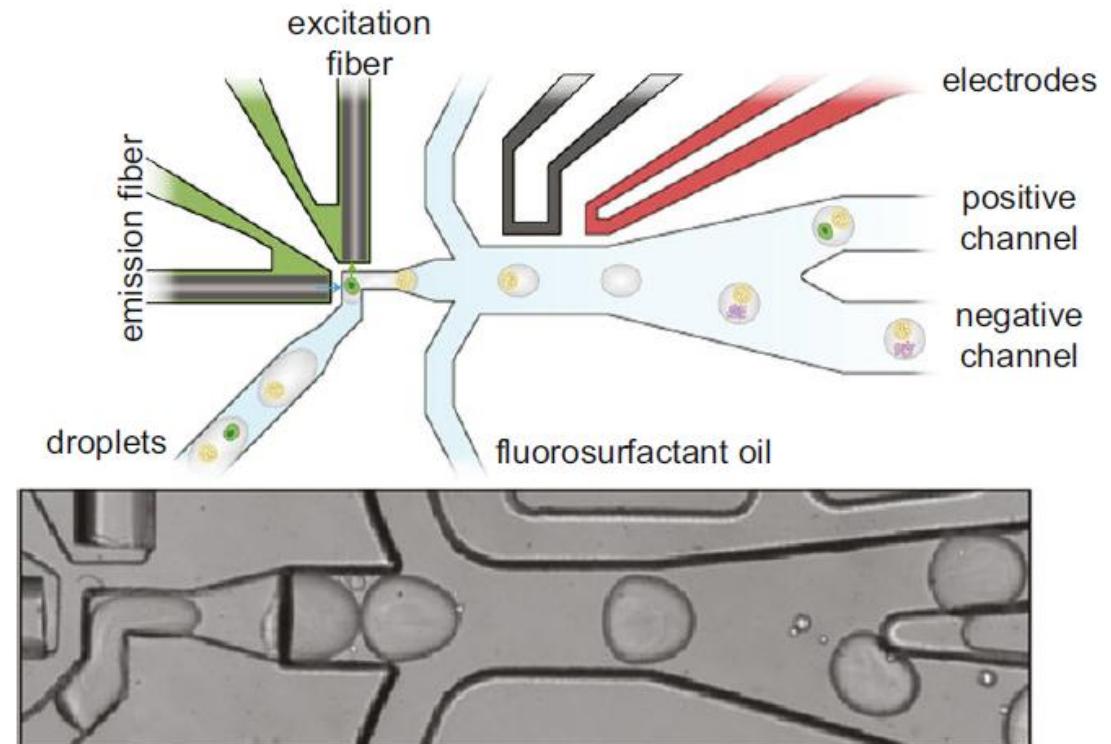
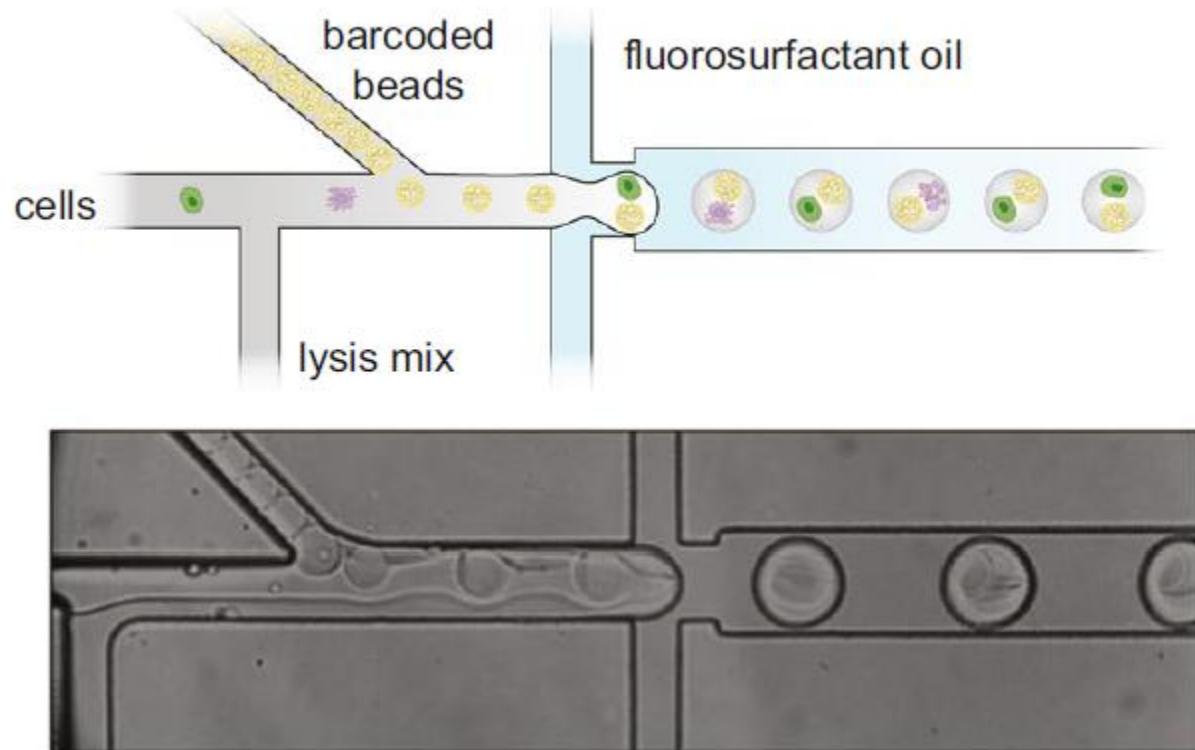
Published online: 08 August 2023

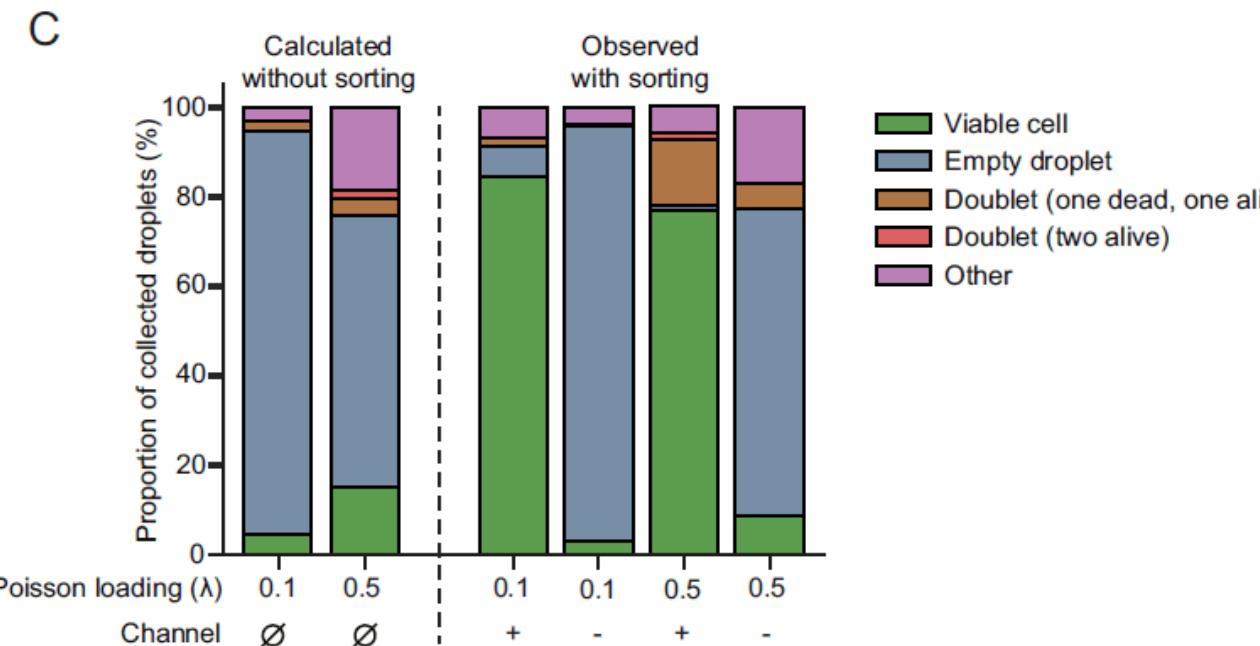
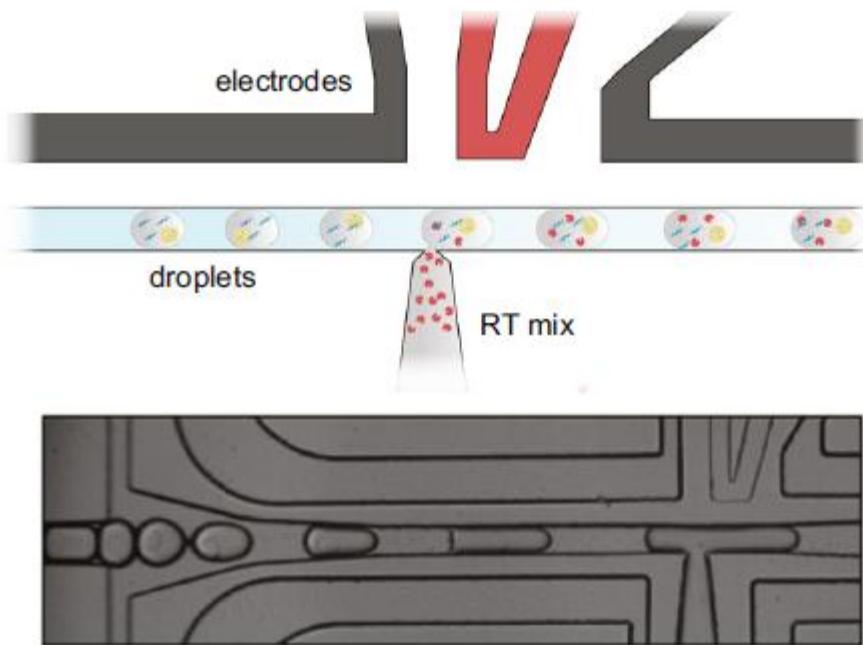


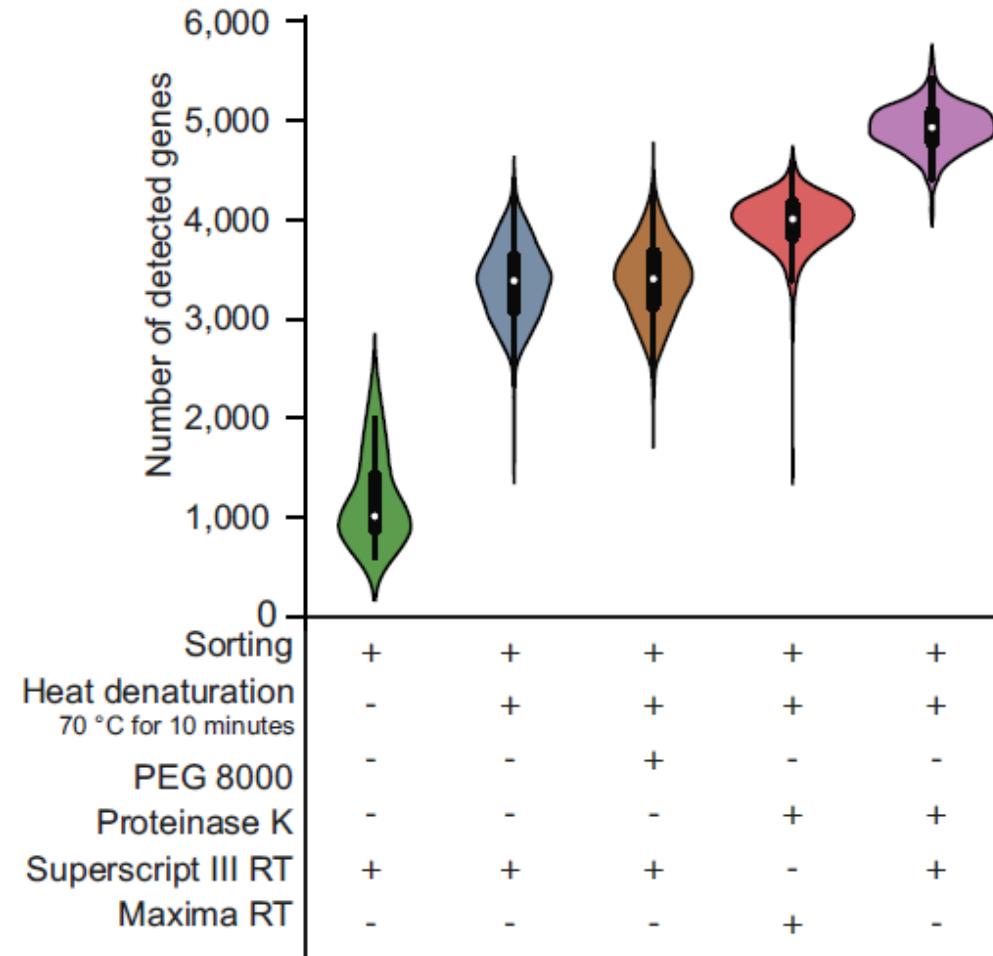
Check for updates

Joachim De Jonghe  ^{1,2,10}, Tomasz S. Kaminski  ^{1,3,10}, David B. Morse ⁴,
Marcin Tabaka ^{5,6}, Anna L. Ellermann  ¹, Timo N. Kohler  ¹, Gianluca Amadei ⁷,
Charlotte E. Handford ⁷, Gregory M. Findlay  ², Magdalena Zernicka-Goetz  ^{7,8},
Sarah A. Teichmann  ⁹ & Florian Hollfelder  ¹ 

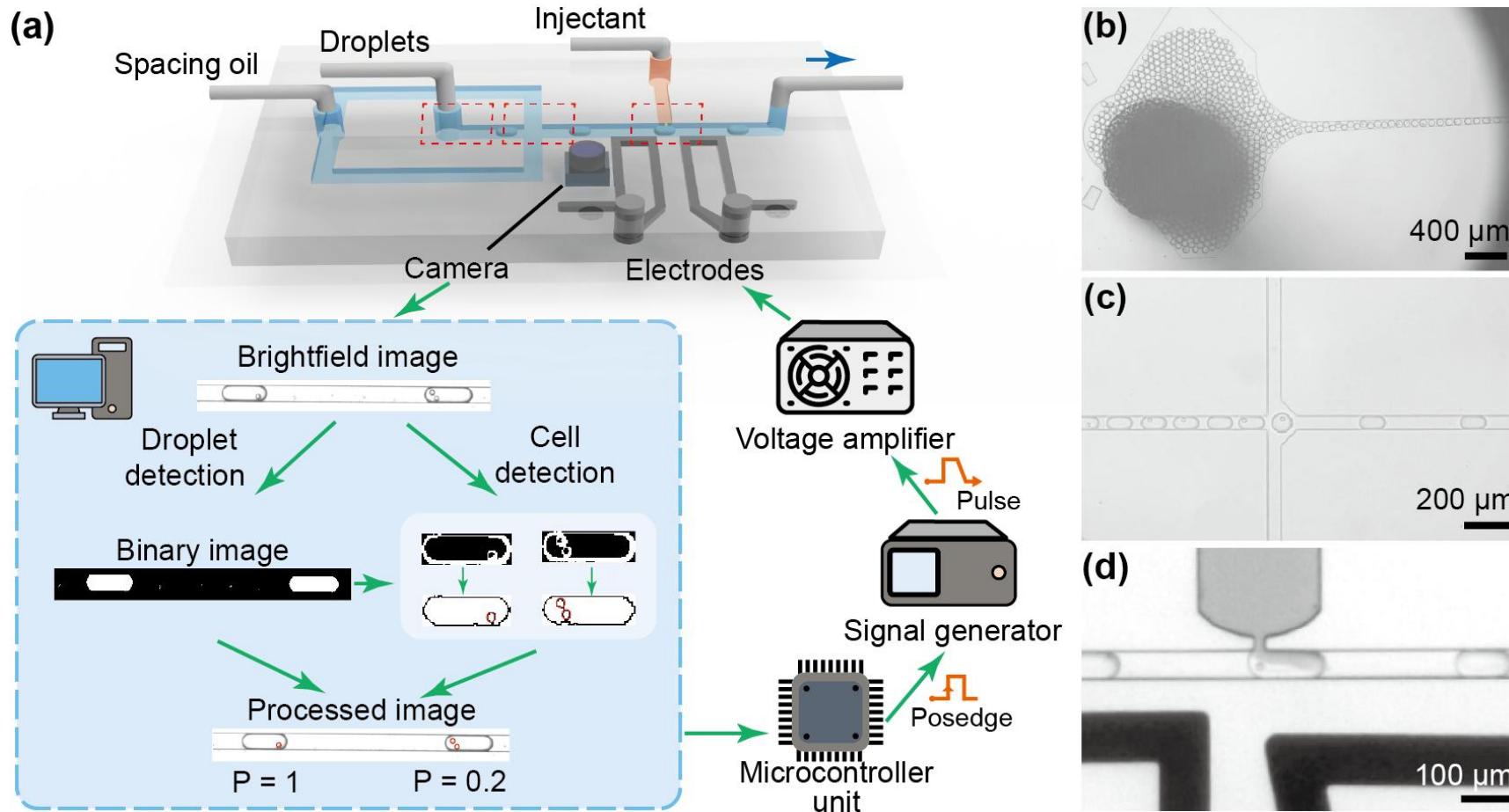
SpinDrop



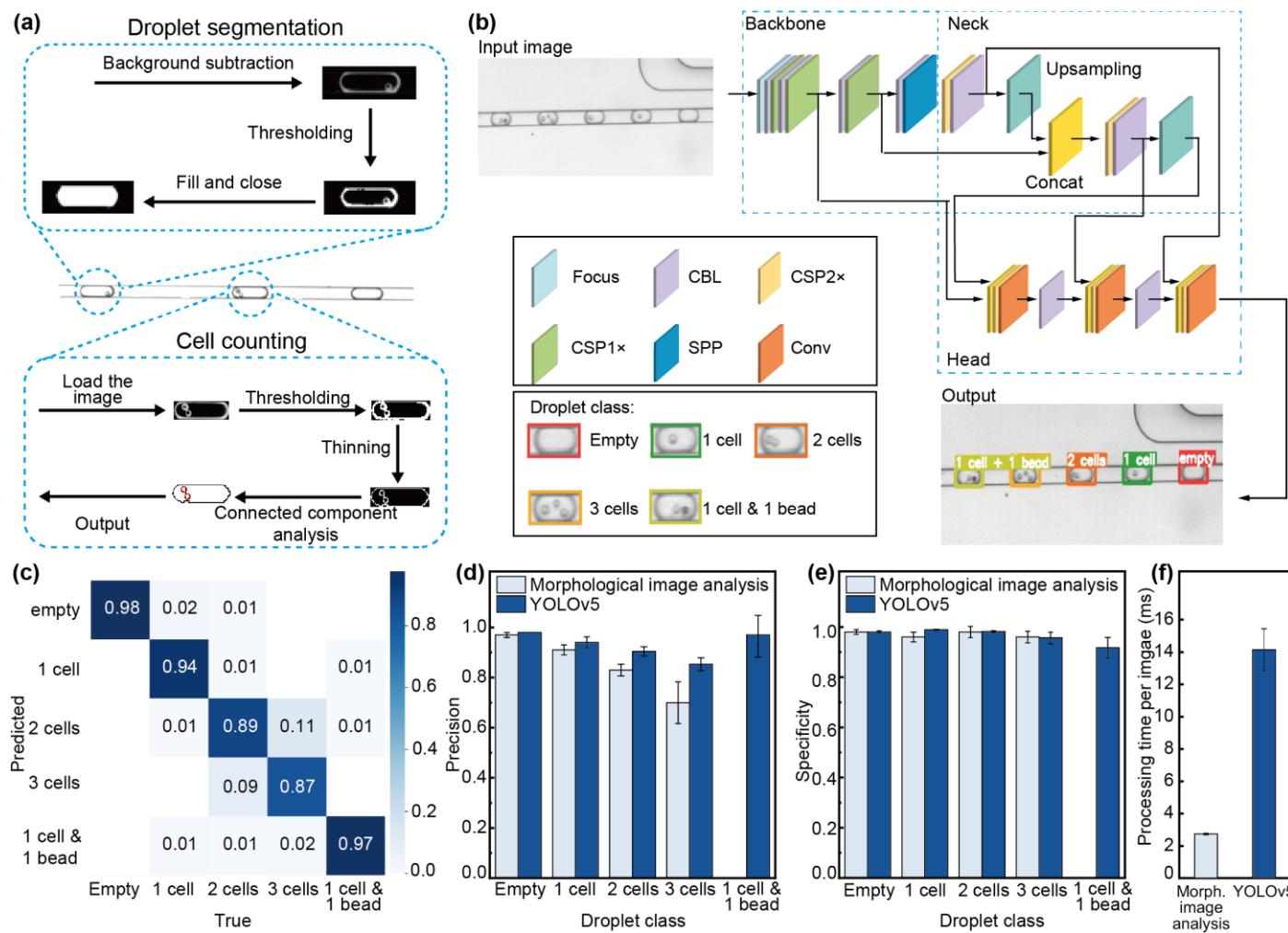


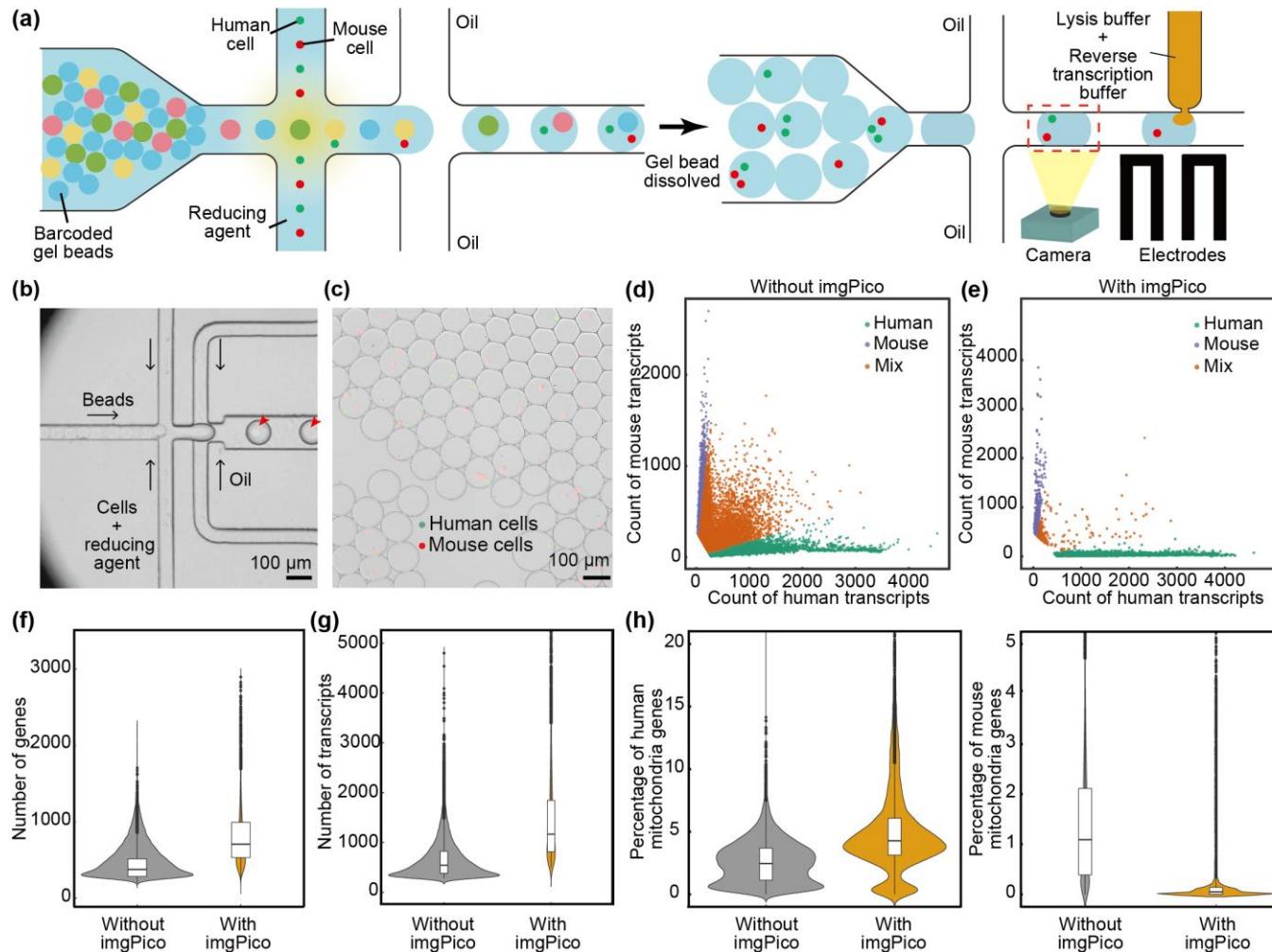


ImgPico: Image-activated pico-injection

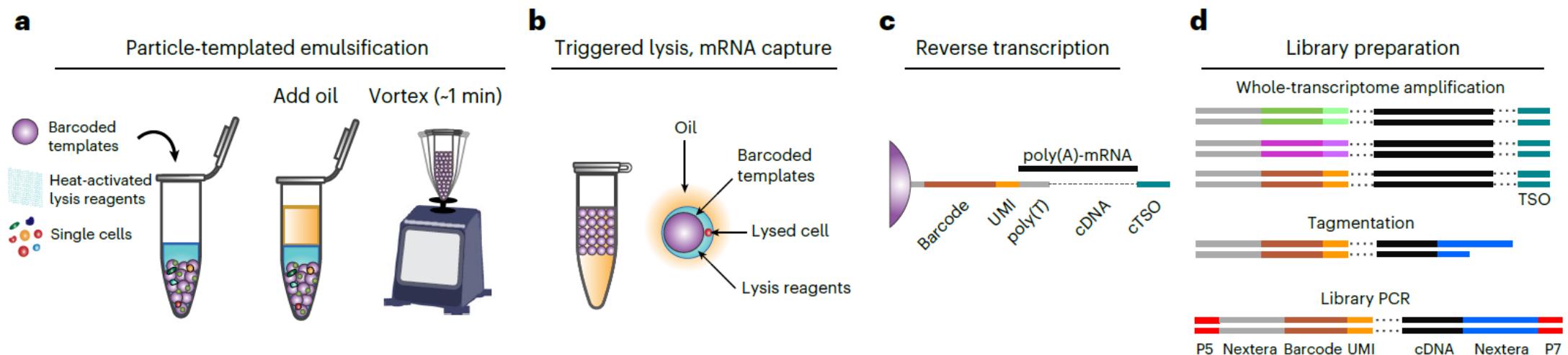


Real-time image analysis of droplets





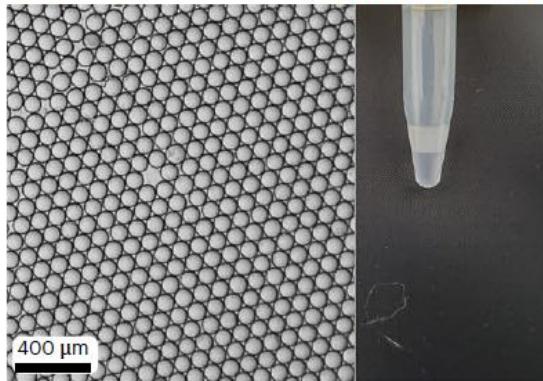
PIP-seq



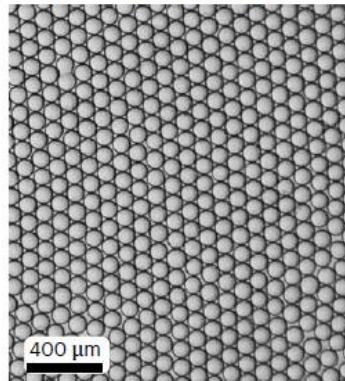
e

Cell number scalability

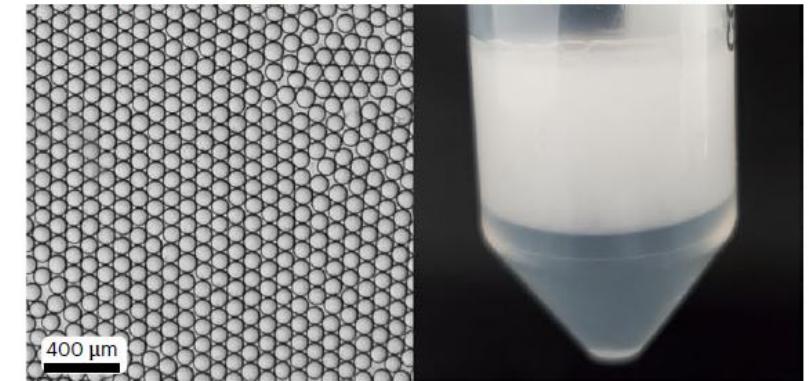
500-ul tube (35 ul PIPs ~4,000 cells cell input)



15-ml Falcon (2 ml PIPs ~250,000 cells cell input)

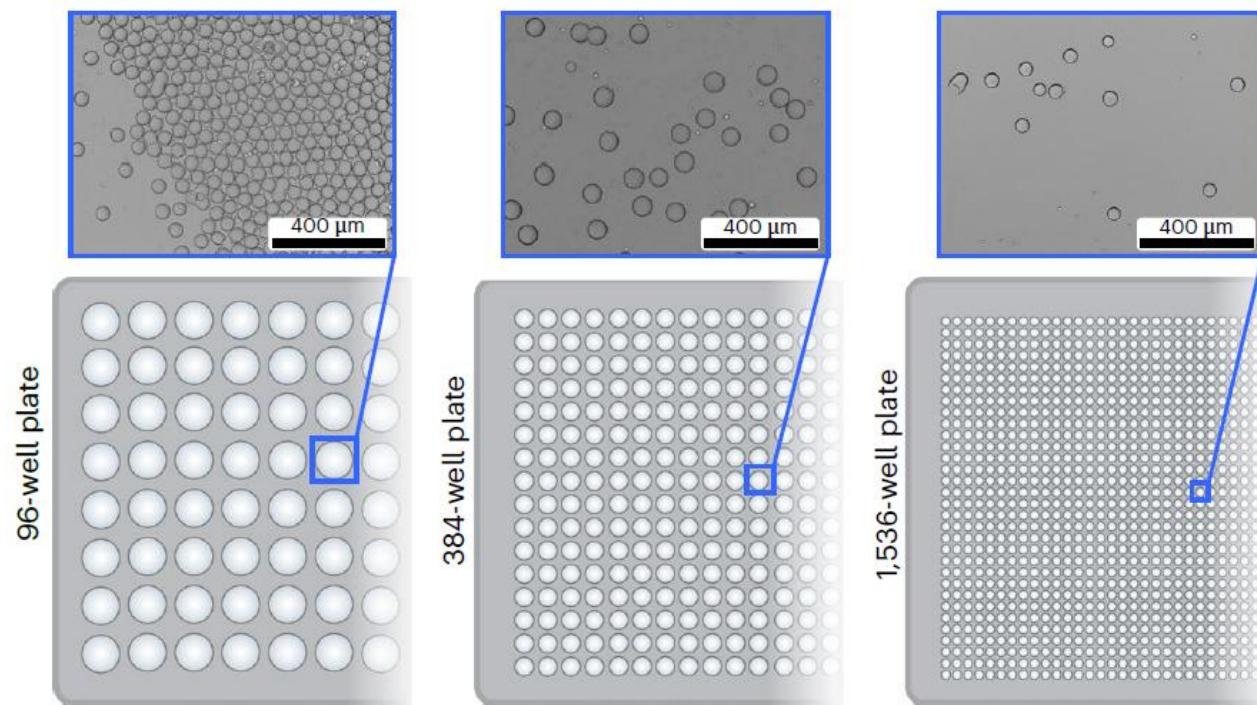


50-ml Falcon (10 ml PIPs ~1 million cells cell input)

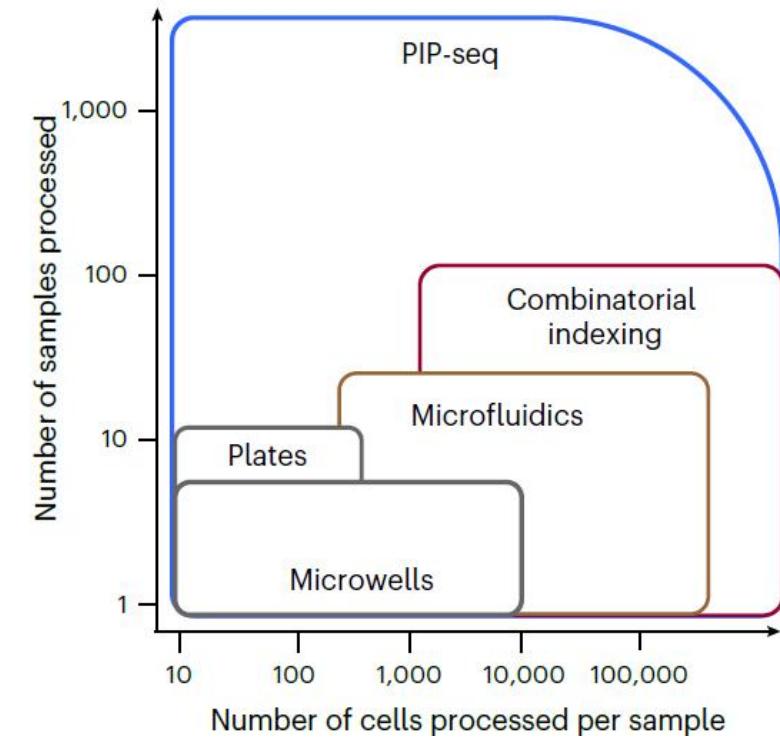


f

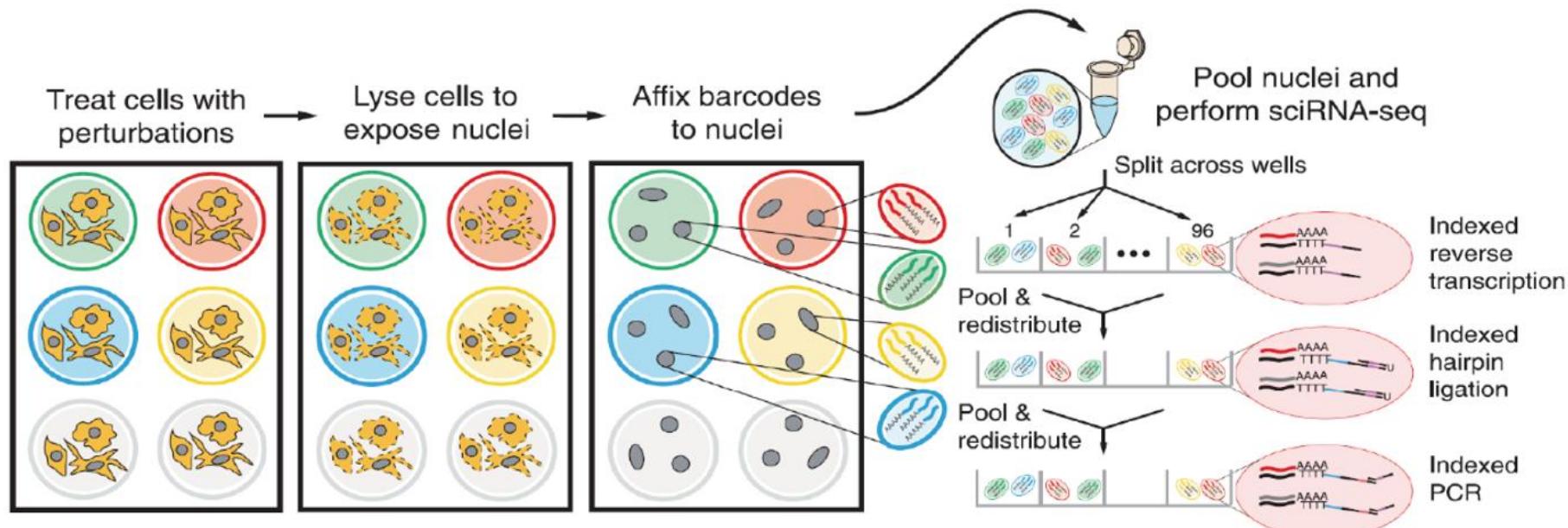
Sample number scalability

**g**

scRNA-seq scalability



CP-seq: combinatorial perturbation sequencing



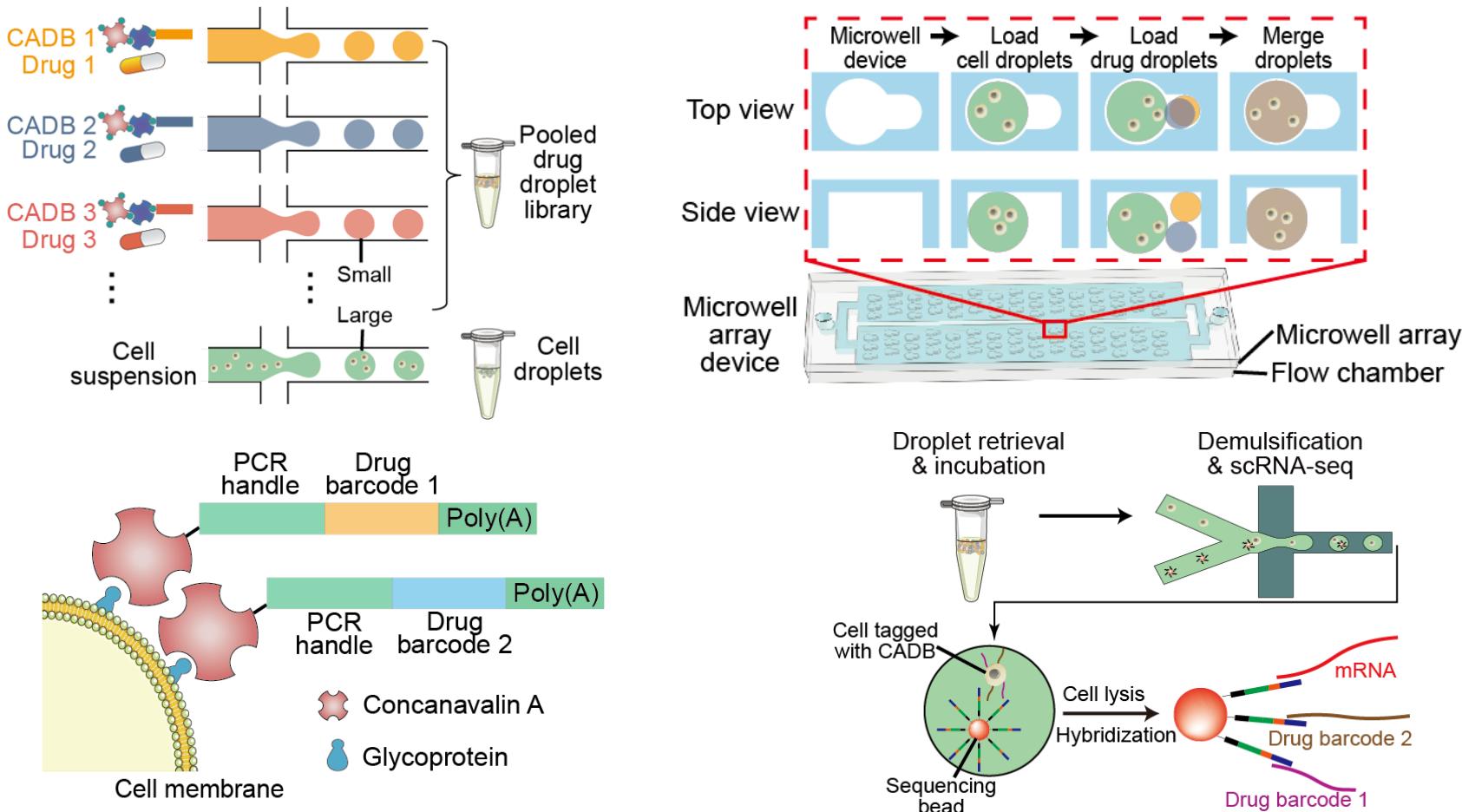
实验条件数目~孔板数目

二重药物组合: $\binom{10^2}{2} = \sim 5 \times 10^3$

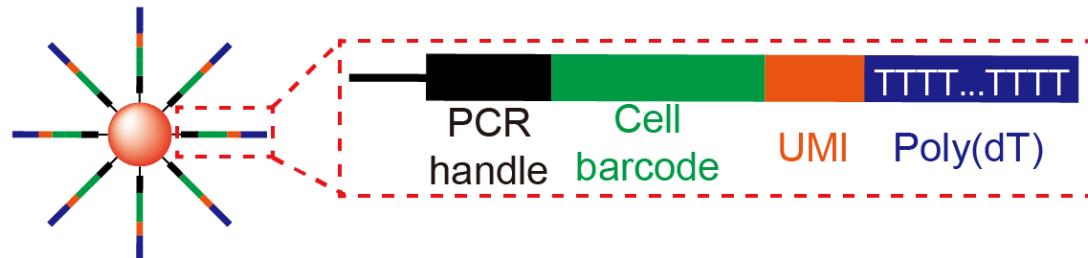
$\binom{10^3}{2} = \sim 5 \times 10^5$

$\binom{10^4}{2} = \sim 5 \times 10^7$

Droplet random pairing + cell membrane labeling = a new method for screening combination drugs



(a)
Sequencing bead

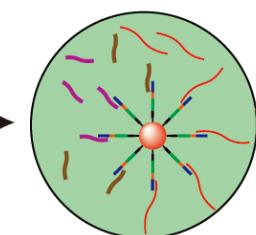


(b)
Biotinylated oligonucleotide

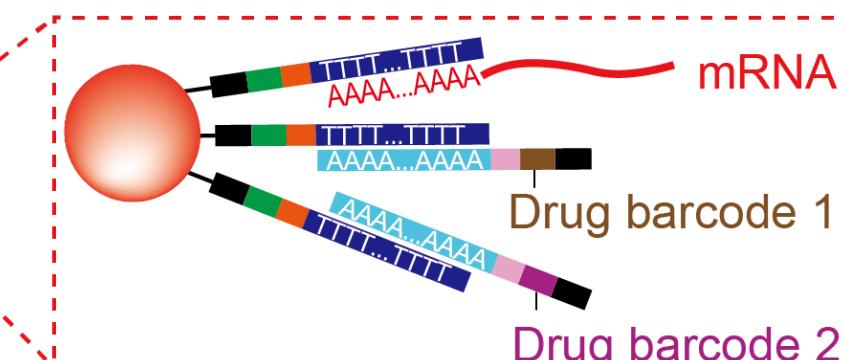


(c)
Cell tagged with combinatorial drug barcodes

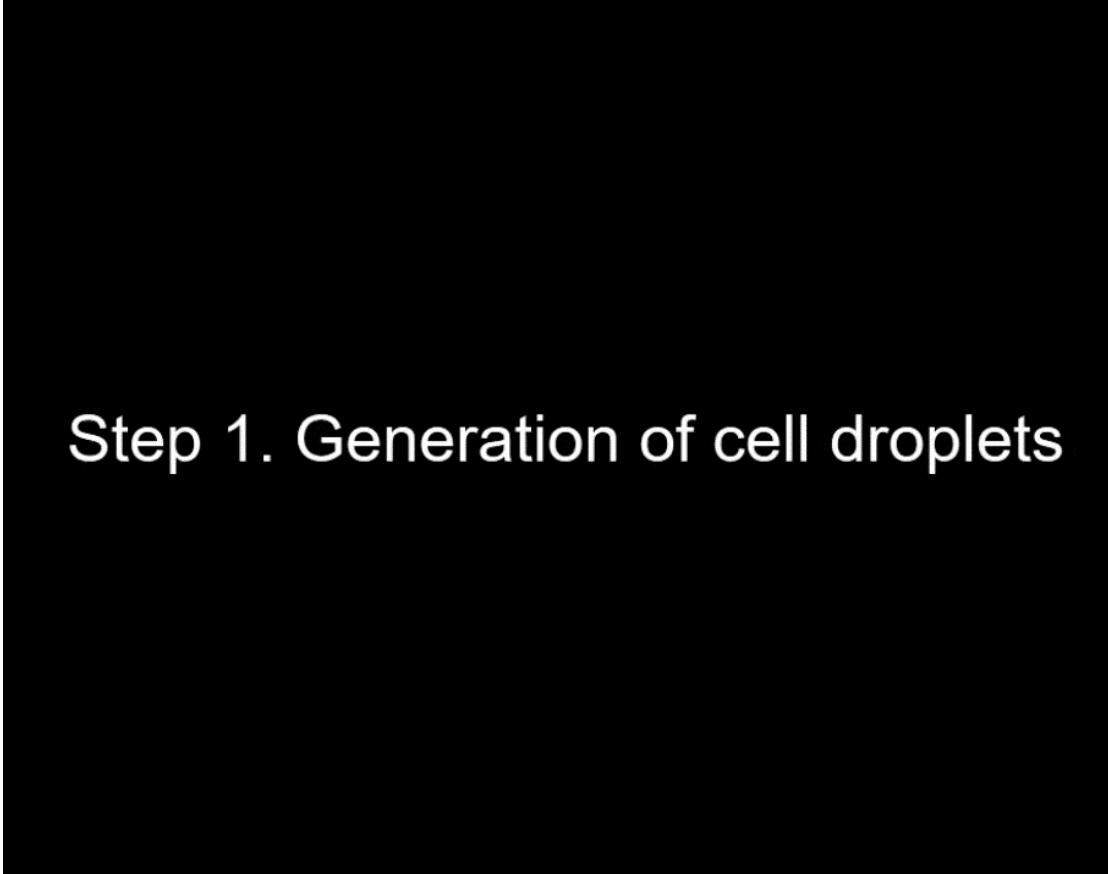
Cell lysis
Hybridization



Sequencing bead after hybridization

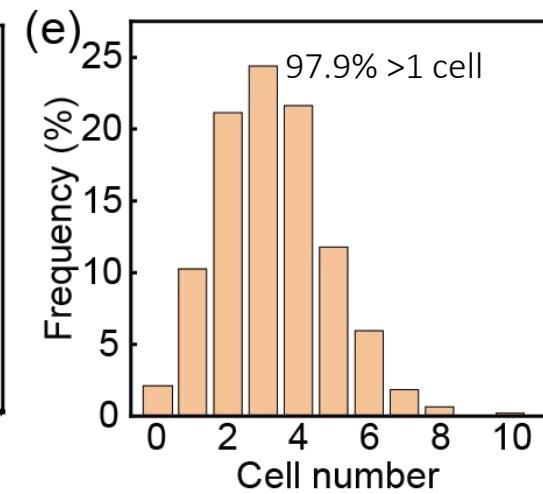
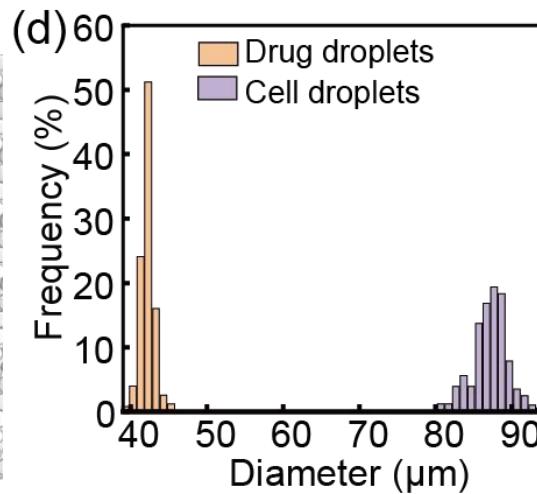
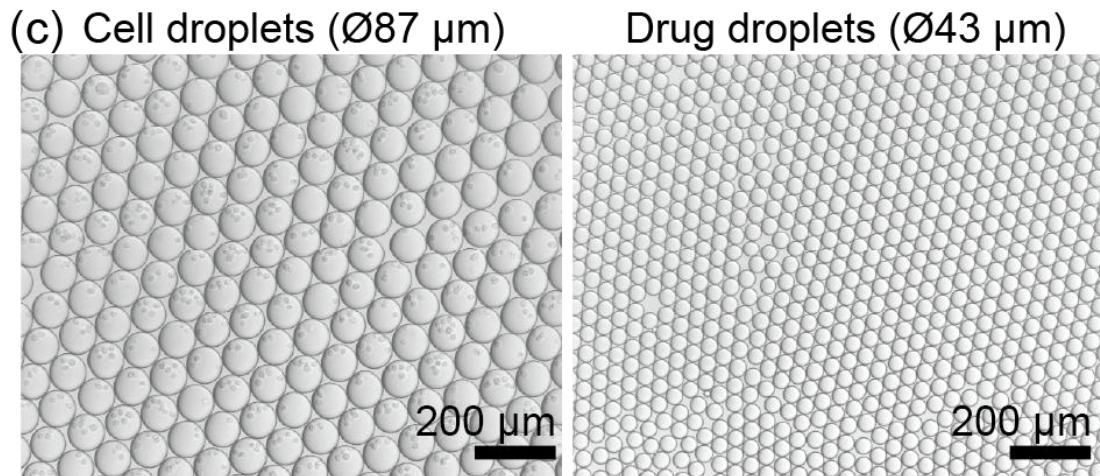
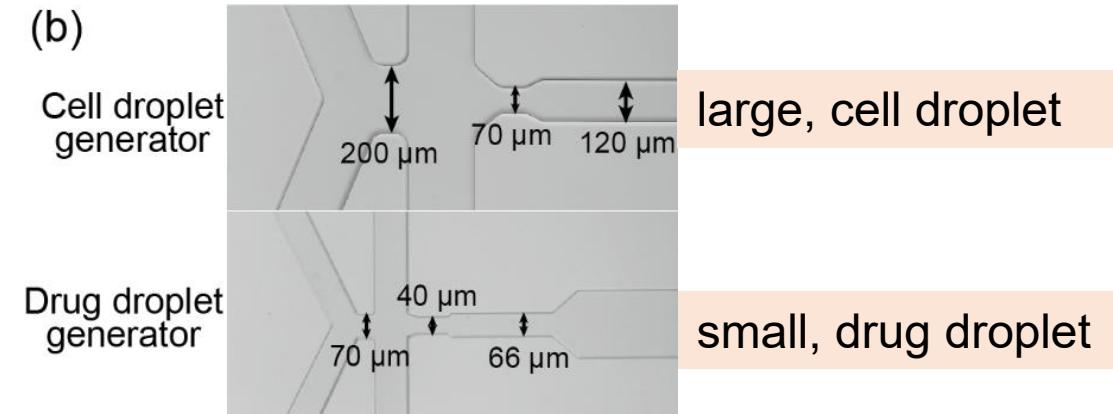
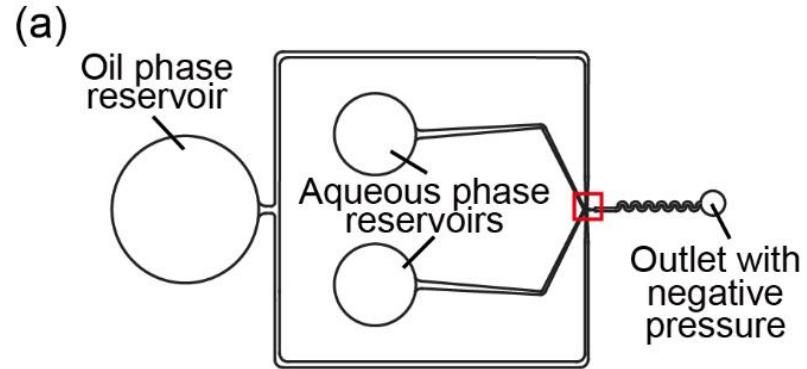


Video demo

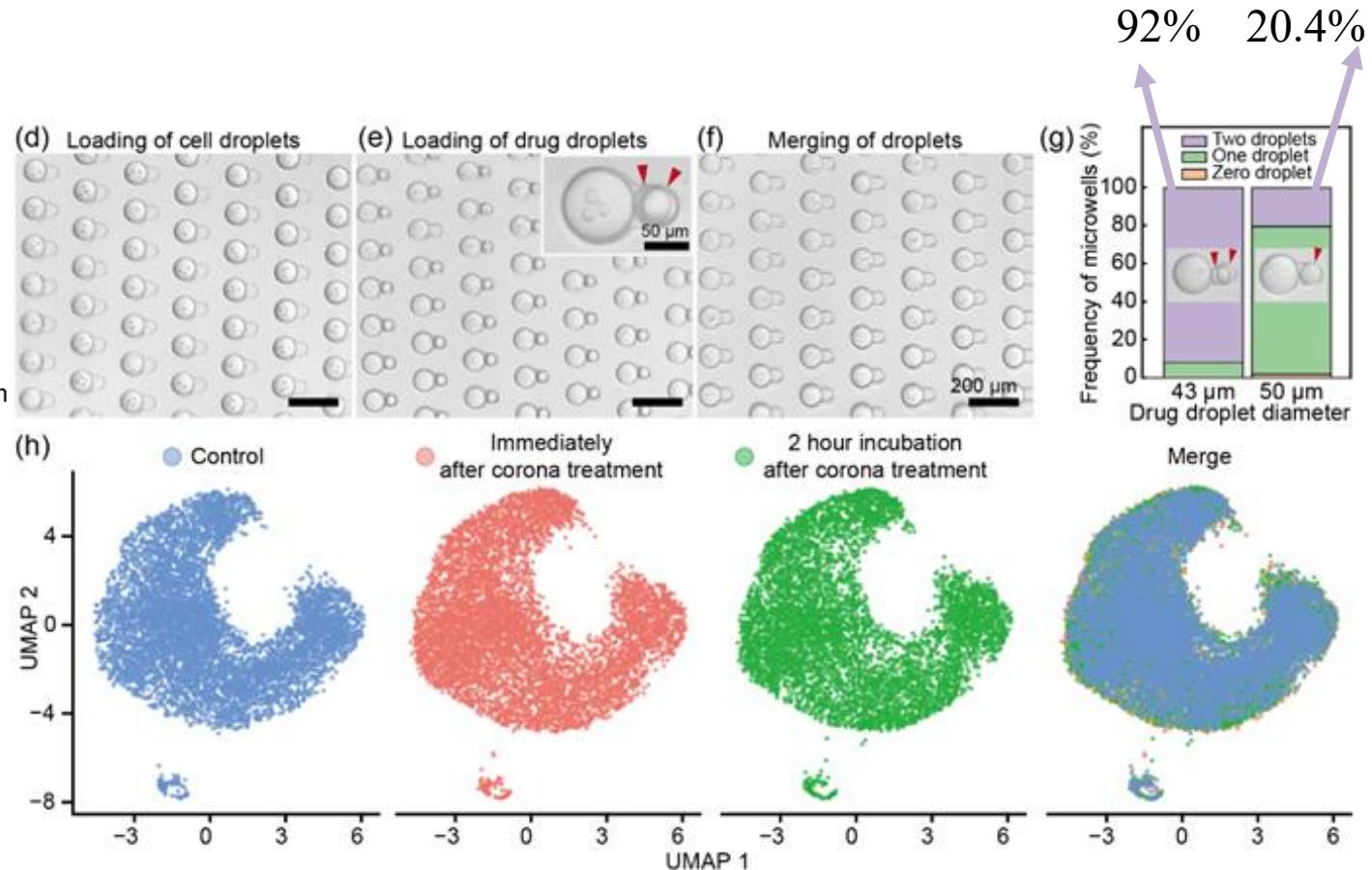
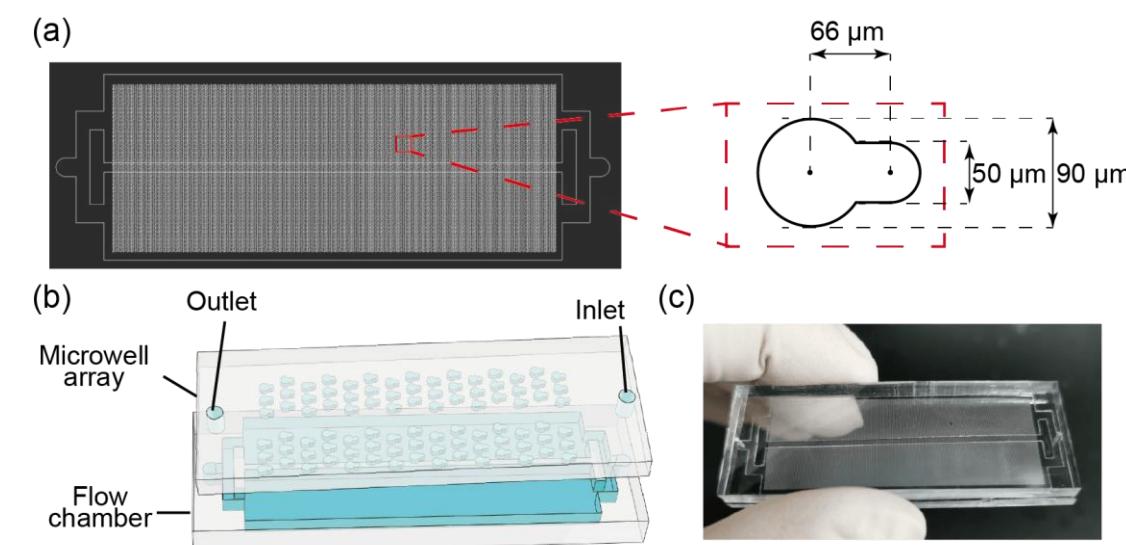


Step 1. Generation of cell droplets

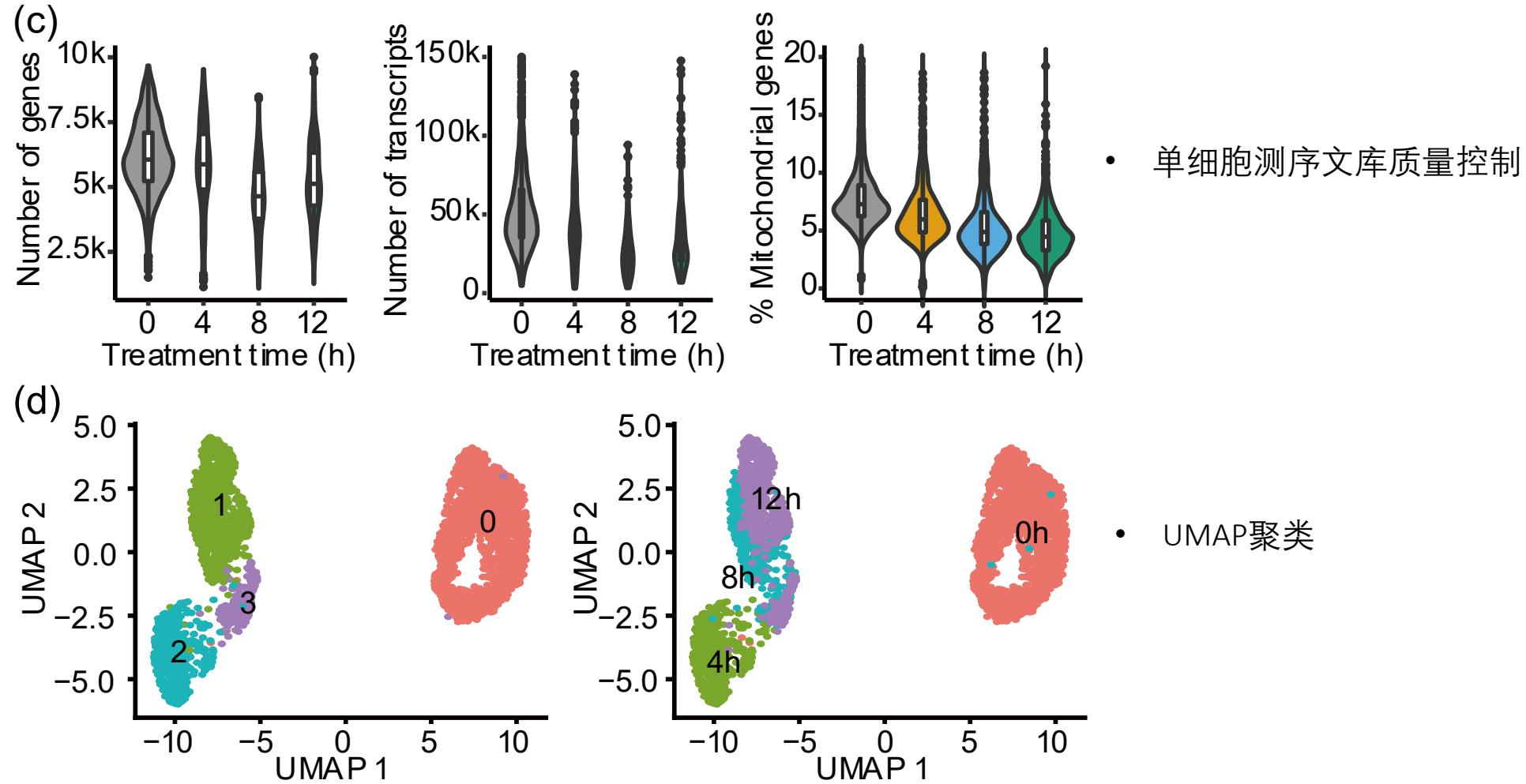
Droplet generation chip and droplet size distribution: The droplet size is controllable and has good stability



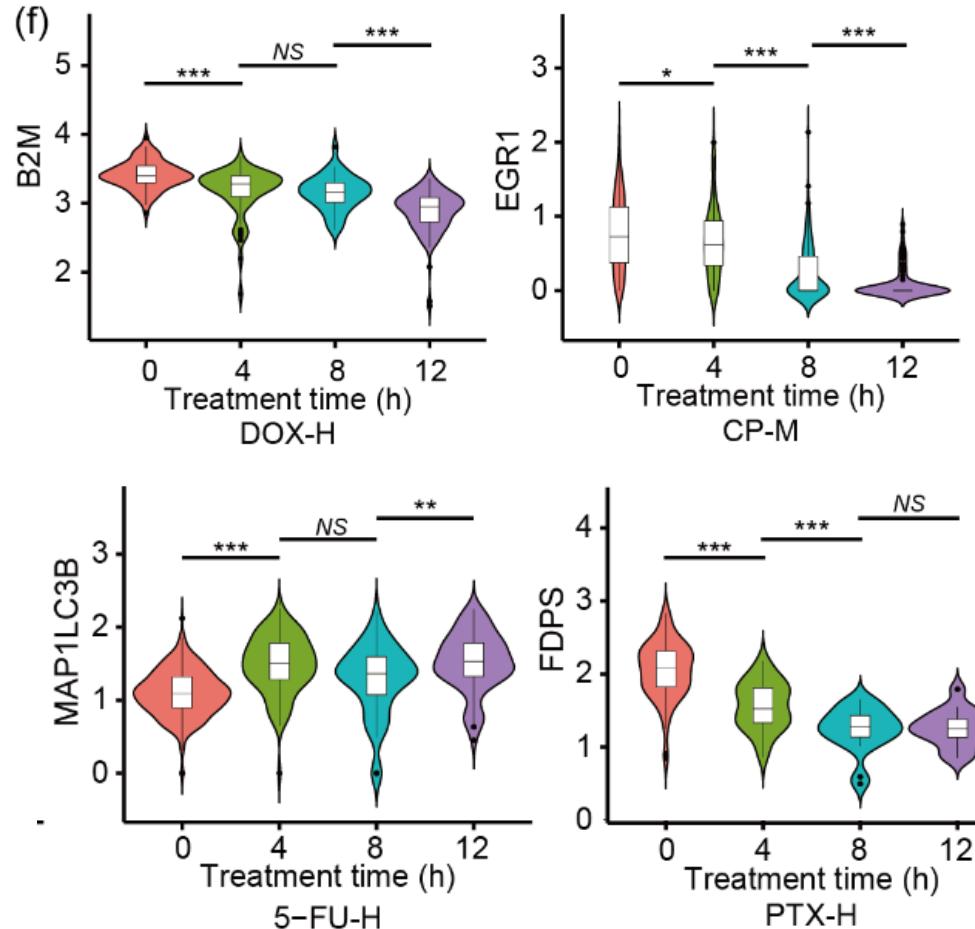
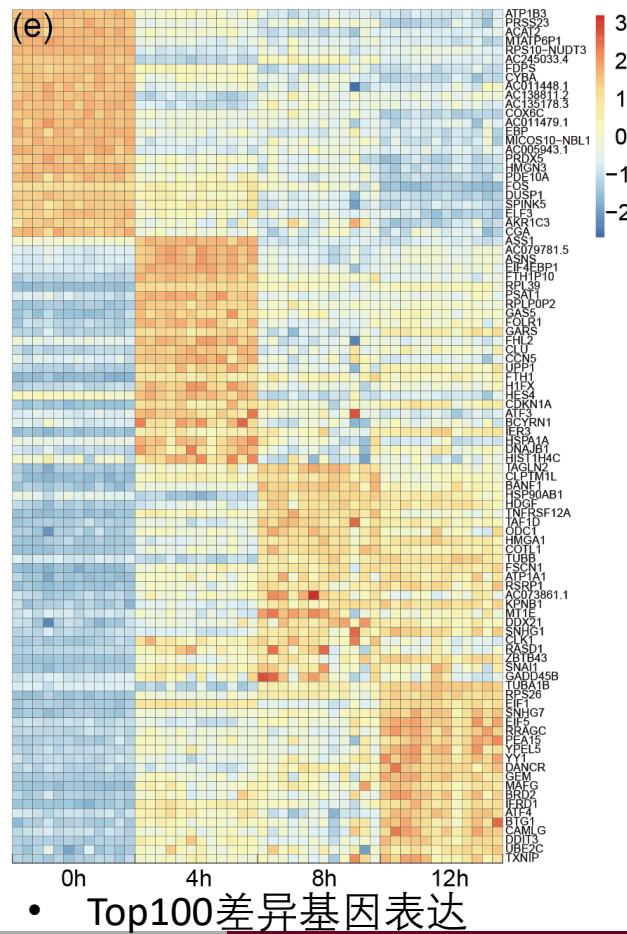
Microporous array chips enable droplet capture, pairing, and fusion: High efficiency droplet capture and fusion, the effect of corona treatment on cell



Validation experiments for single drug treatment



Validation experiments for single drug treatment

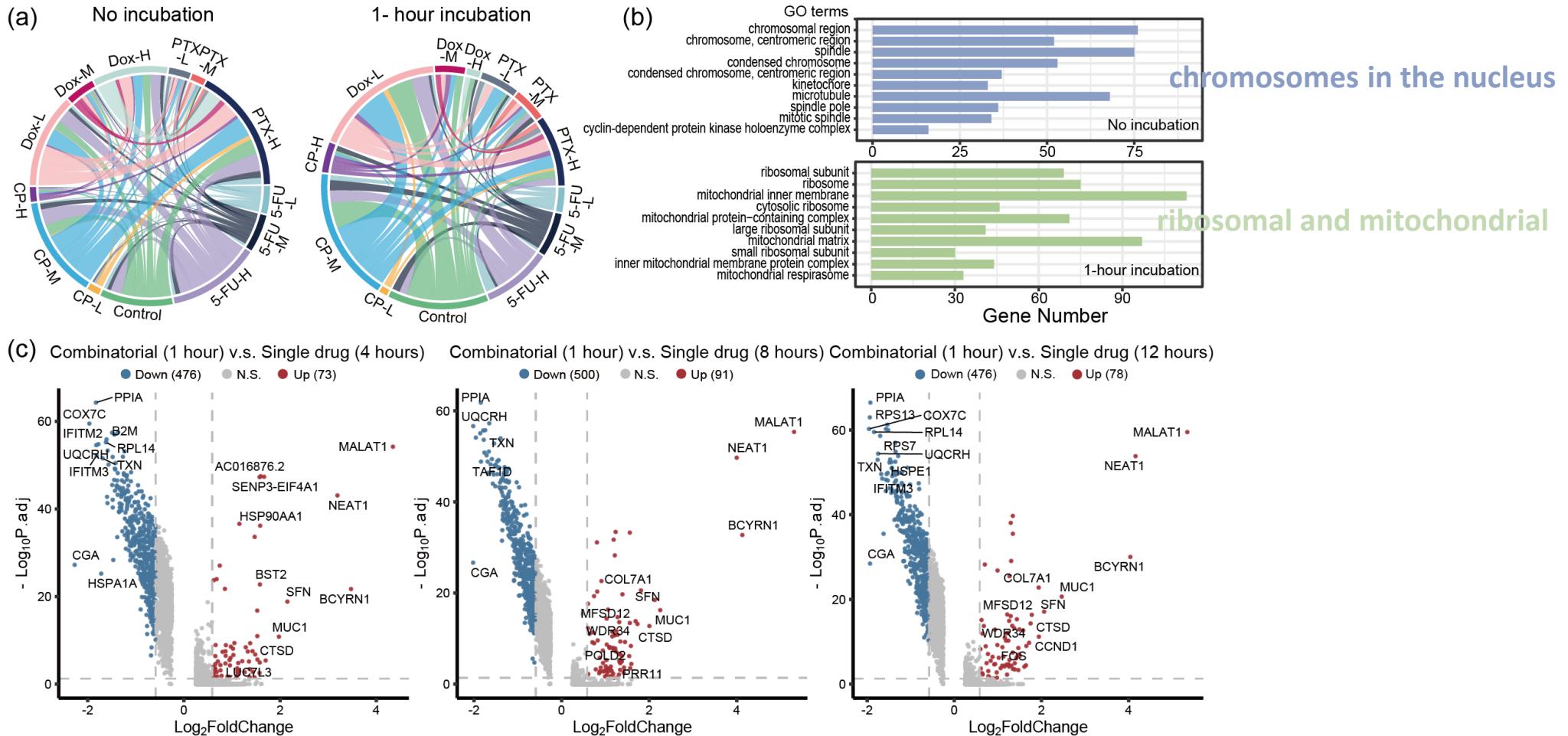


- Top100 差异基因表达

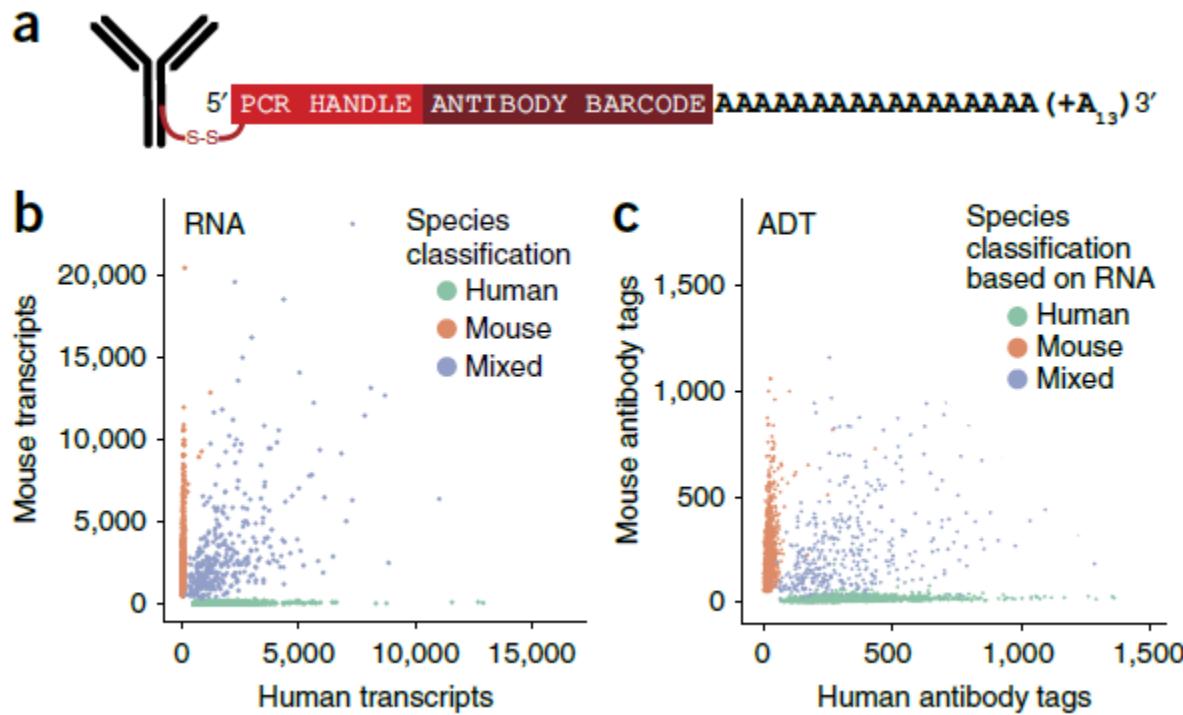
- 与研究报道变化一致的target gene

Validation experiments for combination drug treatment

- GO cellular component ontology analysis



CITE-seq



Associating growth factor secretions and transcriptomes of single cells in nanovials using SEC-seq

Received: 3 April 2023

Accepted: 31 October 2023

Published online: 11 December 2023

Shreya Udani  ^{1,9}, Justin Langerman ^{2,9}, Doyeon Koo  ¹, Sevana Bagdasarian ³, Brian Cheng  ¹, Simran Kang ¹, Citradewi Soemardy ¹, Joseph de Rutte ⁴, Kathrin Plath  ^{2,5,6}  & Dino Di Carlo  ^{1,4,5,7,8} 

