



# 功能基因组学

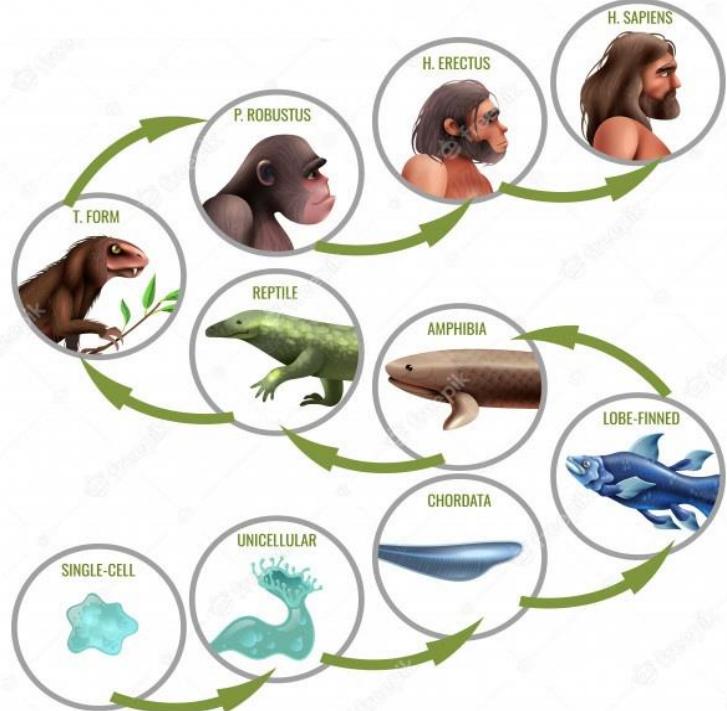
----从细胞水平解析基因组功能

刘龙奇

深圳华大生命科学研究院

2021年11月25日

## 进化



## 发育



- 
- 01** 基因组与细胞命运调控
  - 02** 单细胞组学发展及应用
  - 03** 时空组学发展及应用

01

基因组与细胞命运调控

02

单细胞组学发展及应用

03

时空组学发展及应用

**1944**

STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE  
INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES

INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION  
ISOLATED FROM PNEUMOCOCCUS TYPE III

By OSWALD T. AVERY, M.D., COLIN M. MACLEOD, M.D., AND  
MACLYN McCARTY,\* M.D.

J. Exp. Med. (1944)



Oswald  
Avery

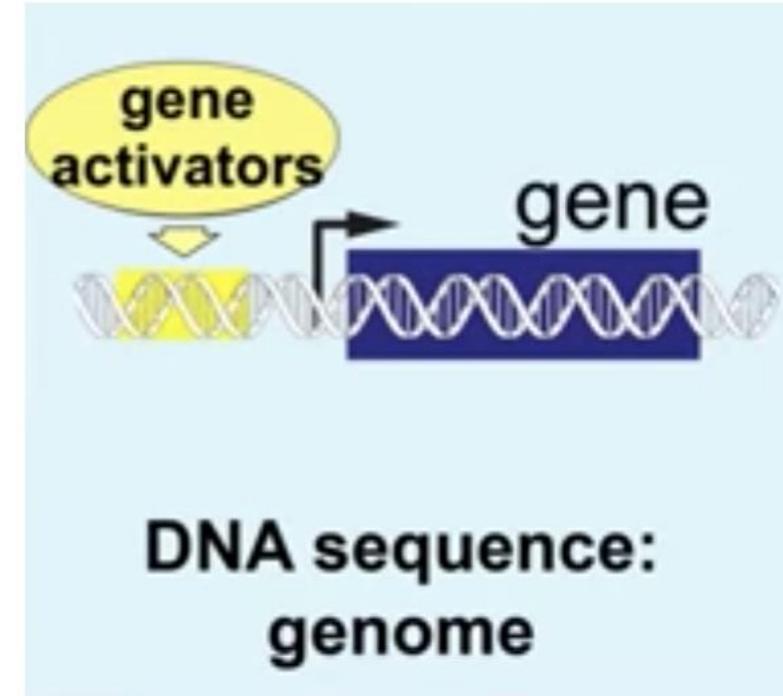


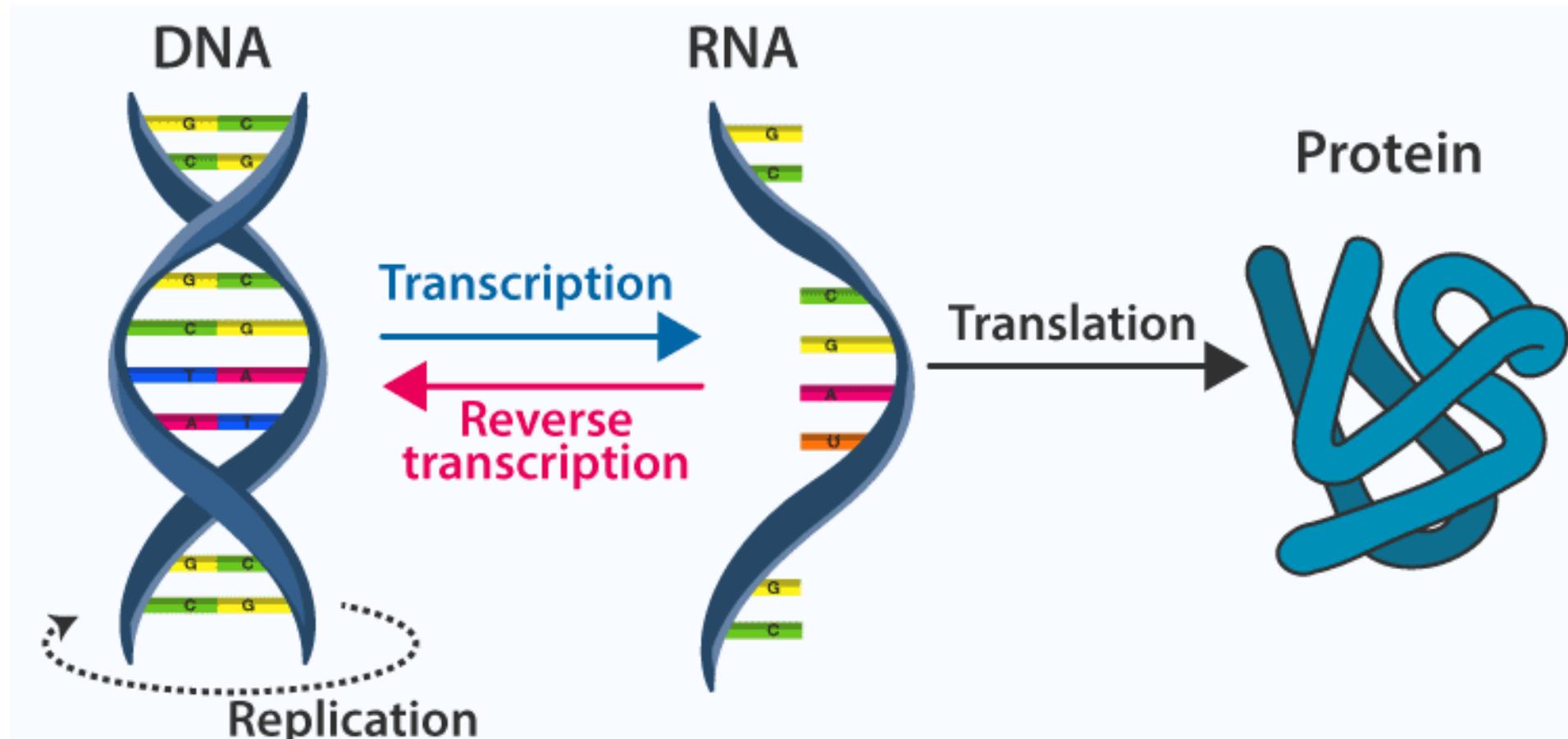
Colin  
MacLeod



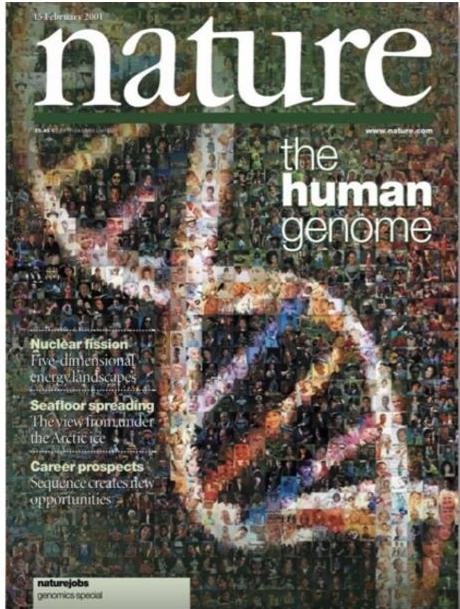
Maclyn  
McCarty

## Genes are made of DNA !





2001

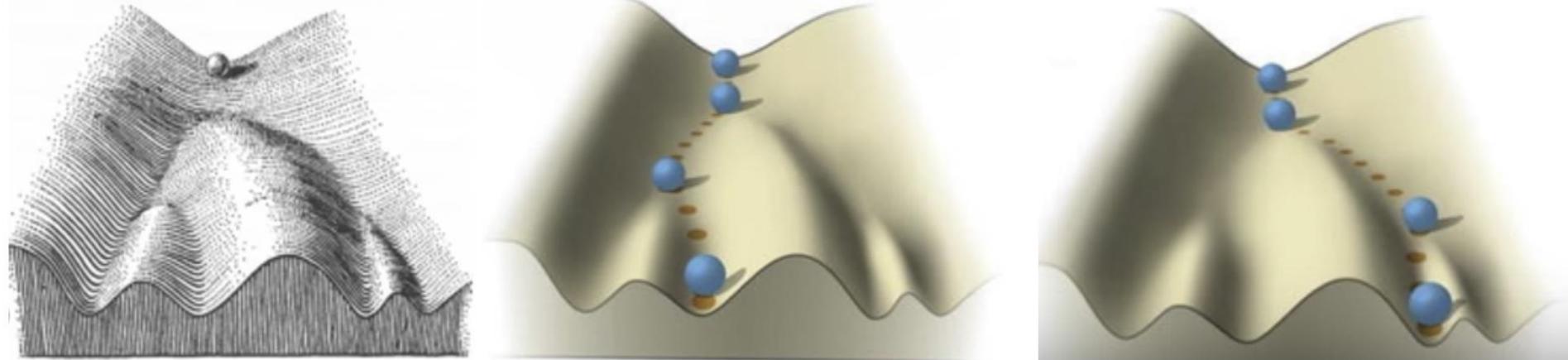


## The logic behind the HGP:

1. Genes determine aging;
2. Genes determine diseases;
3. Genetic analyses provide diagnosis and therapy (personalized medicine)

1. Why so few protein coding genes (~25000) ?
2. Aren't we more complicated than worms, flies and fish?
3. How can past experiences or our environment shape who we are?
4. Why there are so many cell types, while they share the same genome sequences ?

Genes + environment = phenotype



The Waddington landscape (Conrad waddington, 1942)

**Epi-genetics** (Epi: over, above, in addition)



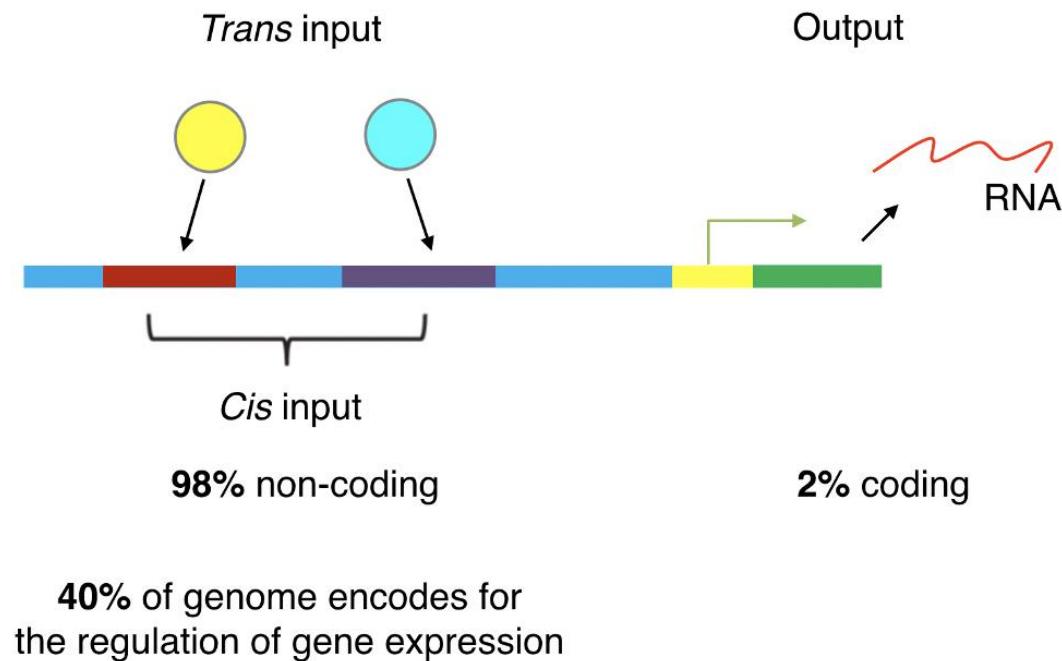
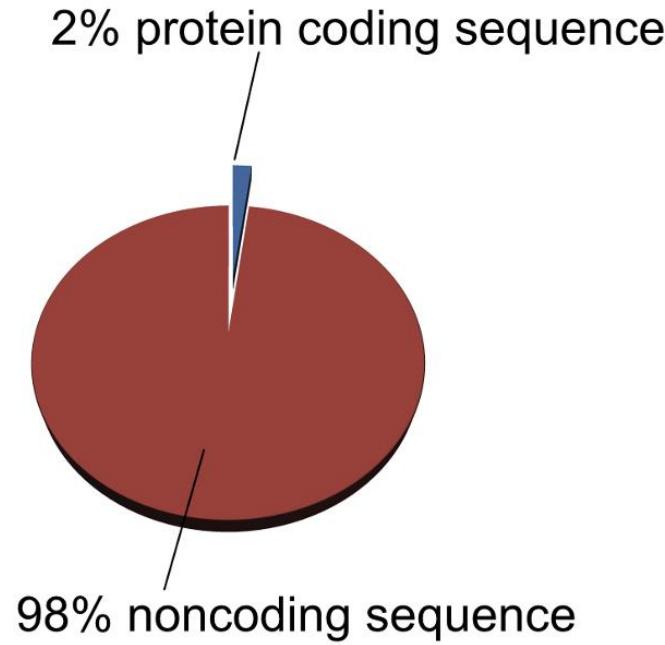
**Genes = Computer hardware  
epigenetics = Computer software**

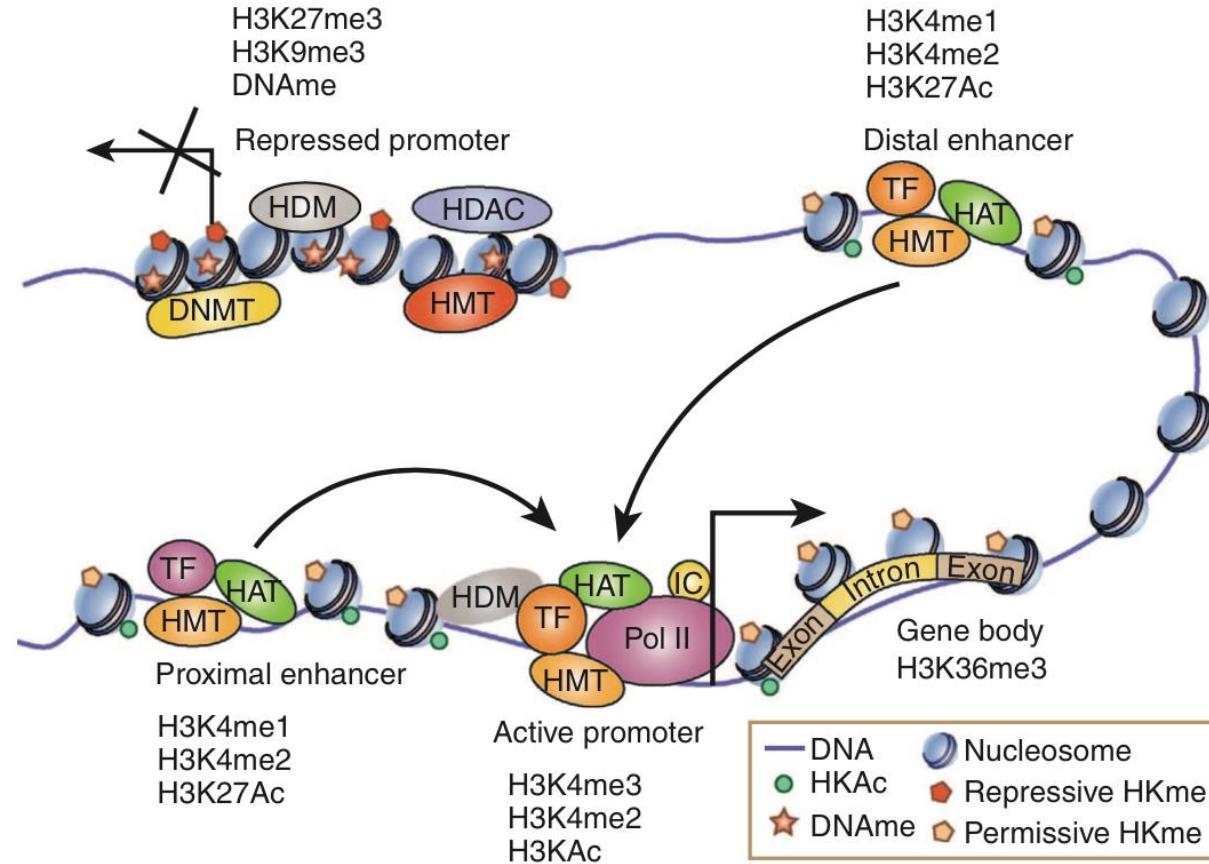


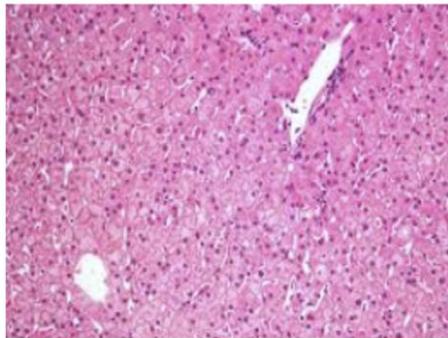
Computer hardware



Computer software





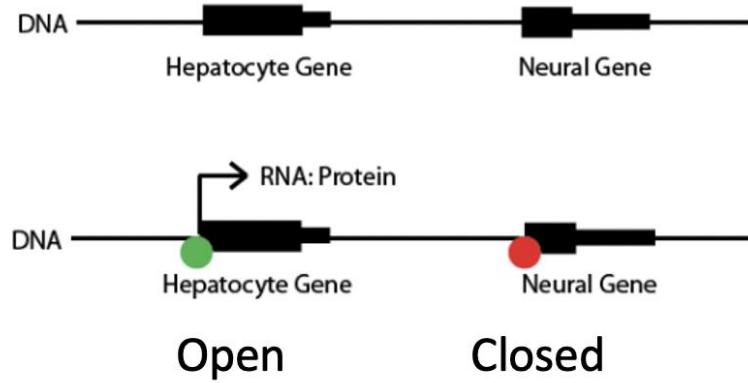


Hepatocytes = liver

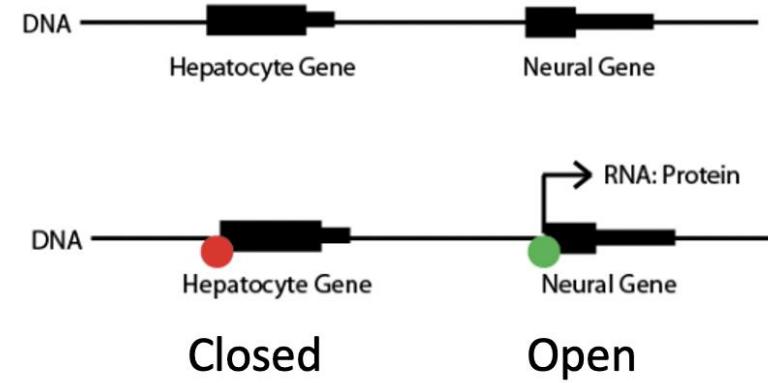


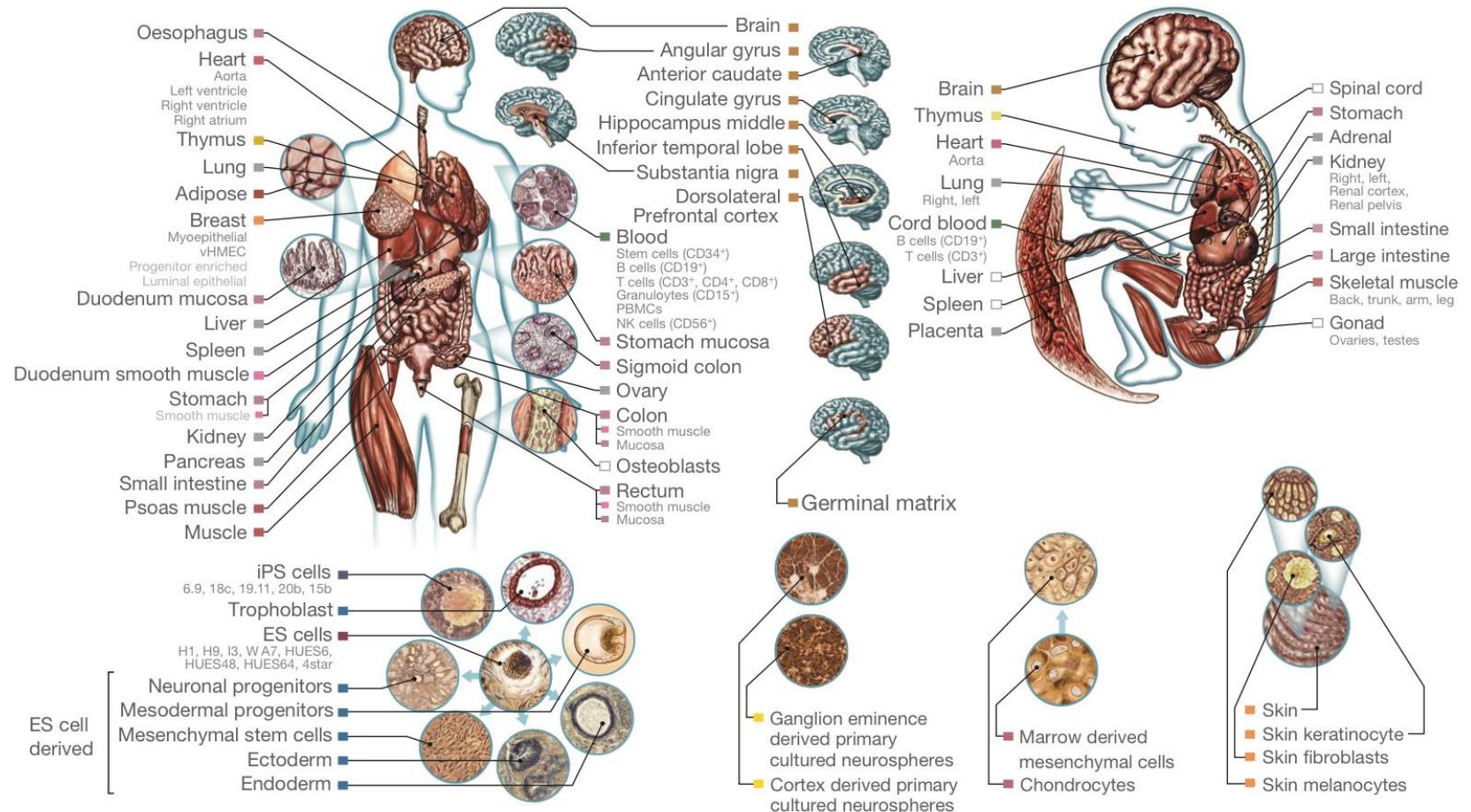
Neurons = brain

break down toxic substances



process information





显微镜



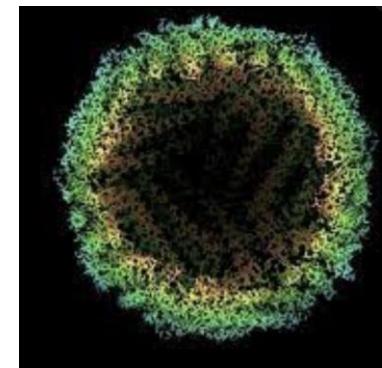
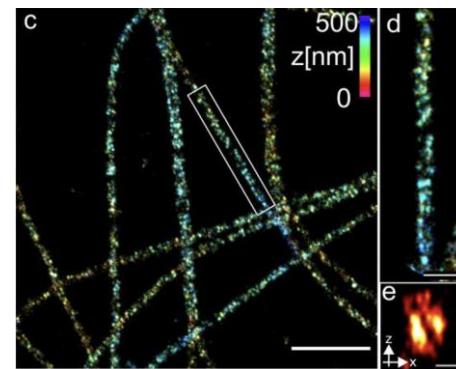
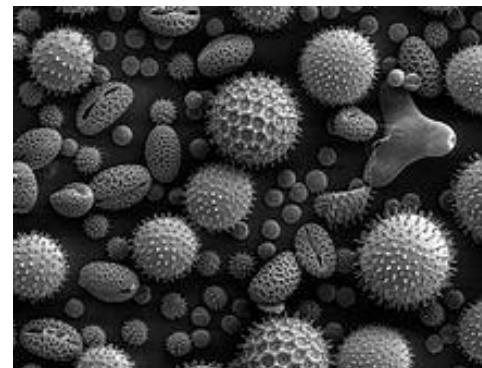
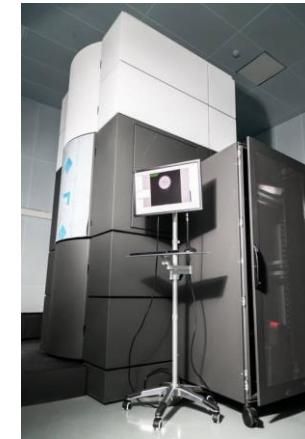
扫描电子显微镜



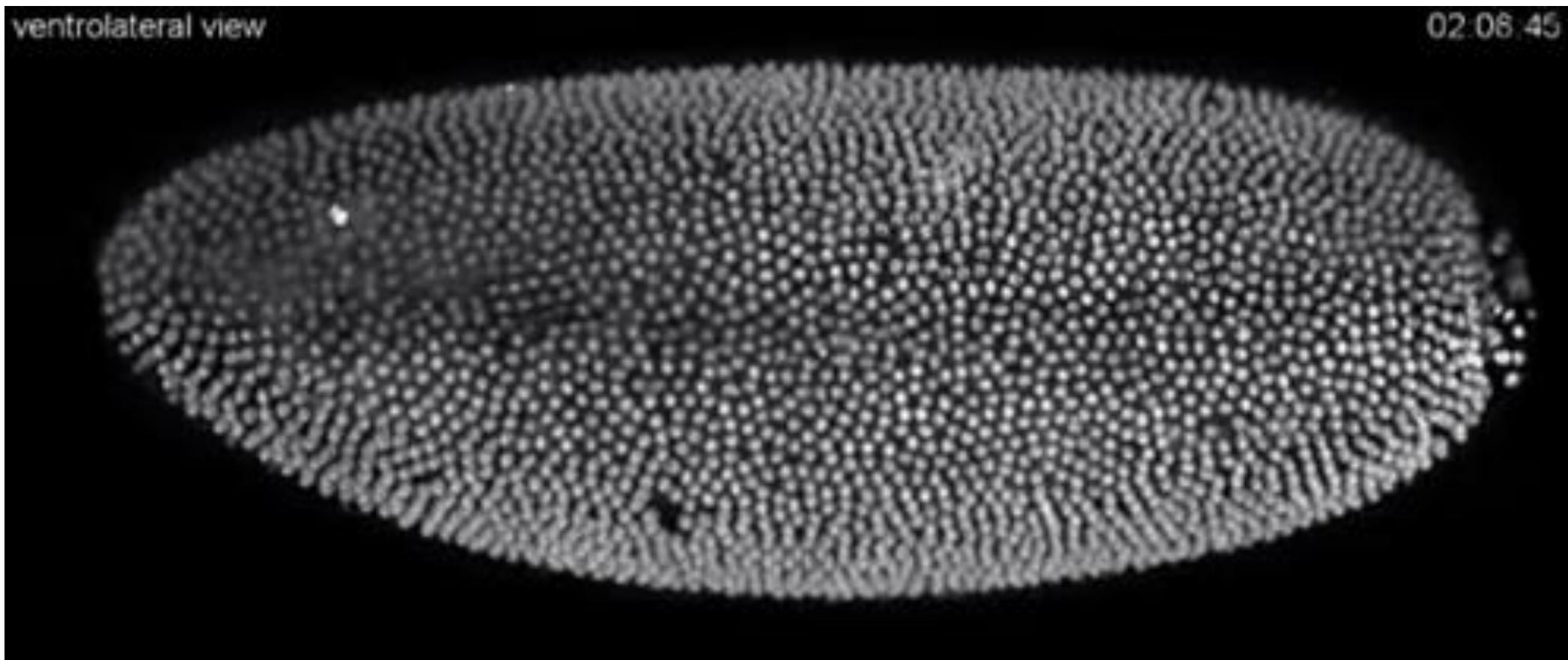
STORM 超高分辨率显微镜



冷冻电镜

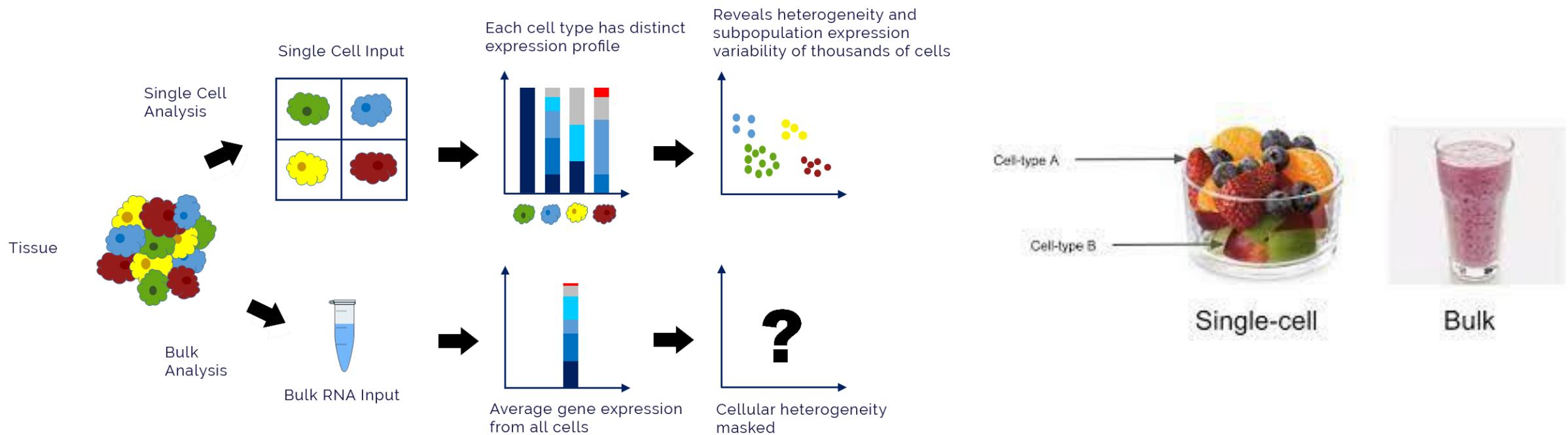


- 技术进步是科学发展的主要动力之一。



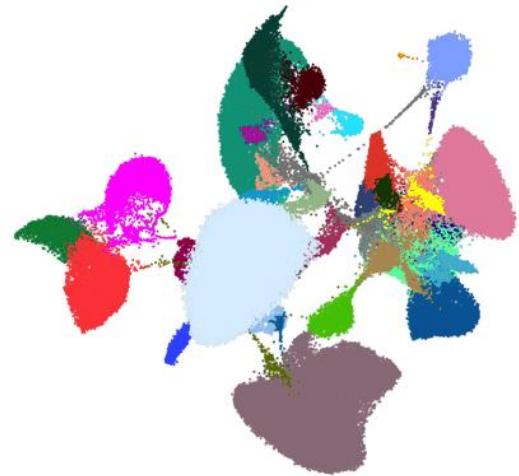
- 单细胞分辨率（单细胞组学）
- 时间-空间信息（时空组学）

- 
- 01** 基因组与细胞命运调控
  - 02** 单细胞组学发展及应用
  - 03** 时空组学发展及应用



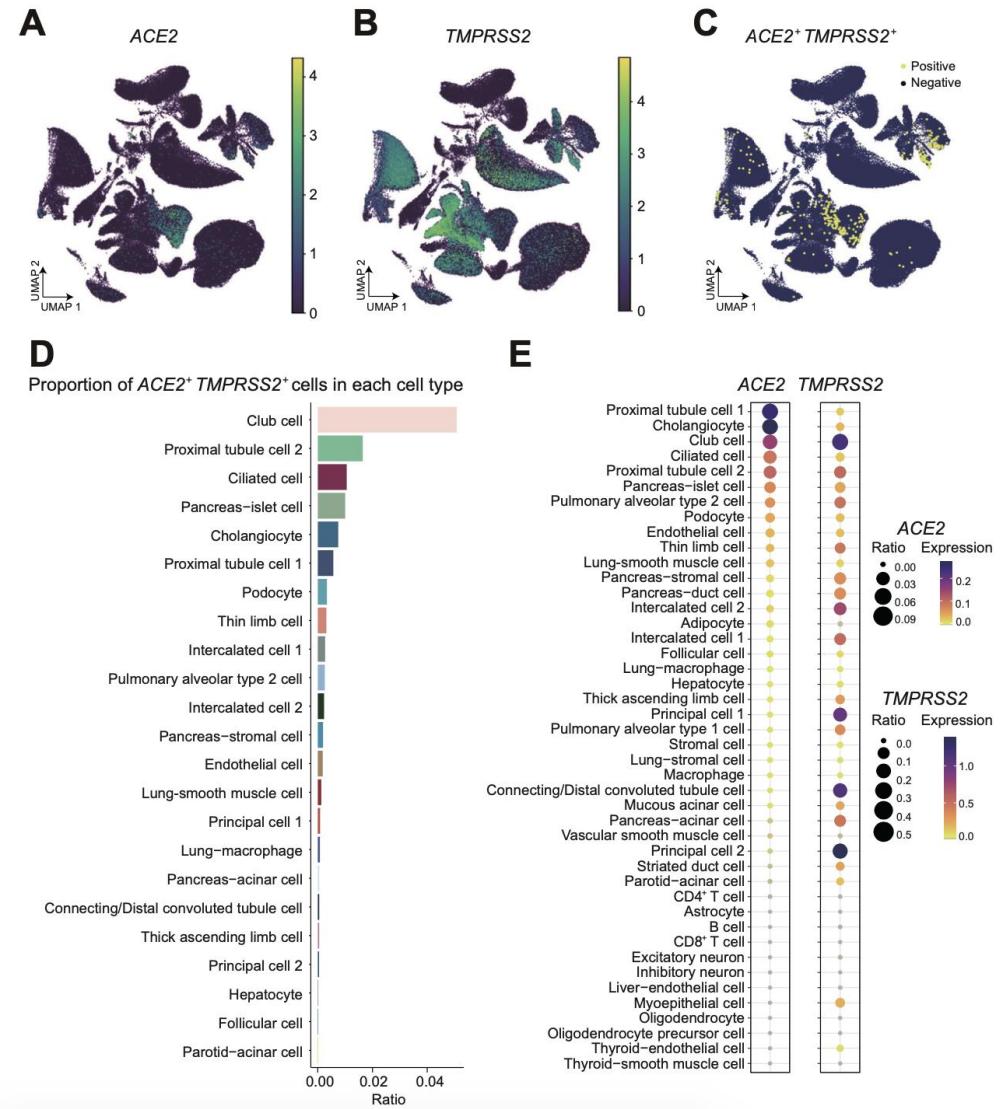
# Single-cell atlas of a non-human primate reveals new pathogenic mechanisms of COVID-19

Lei Han,<sup>1,2,15</sup> Xiaoyu Wei,<sup>1,2,3,15</sup> Chuanyu Liu,<sup>1,2,14,15</sup> Giacomo Volpe,<sup>4,15</sup> Zhifeng Wang,<sup>1,2,15</sup> Taotao Pan,<sup>1,2,3</sup> Yue Yuan,<sup>1,2,3</sup> Ying Lei,<sup>1,2,14</sup> Yiwei Lai,<sup>4</sup> Carl Ward,<sup>4</sup> Yeya Yu,<sup>1,2,5</sup> Mingyue Wang,<sup>1,2,3</sup> Quan Shi,<sup>1,2,6</sup> Tao Wu,<sup>1,2</sup> Liang Wu,<sup>1,2,3</sup> Ya Liu,<sup>1,2</sup> Chunqing Wang,<sup>1,2,3</sup> Yuanhang Zhang,<sup>1,2,3</sup> Haixi Sun,<sup>1,2</sup> Hao Yu,<sup>1,2</sup> Zhenkun Zhuang,<sup>1,2,7</sup> Tingting Tang,<sup>1,2</sup> Yunting Huang,<sup>2</sup> Haorong Lu,<sup>2</sup> Liqin Xu,<sup>1,2,6</sup> Jiangshan Xu,<sup>1,2,3</sup> Mengnan Cheng,<sup>1,2,3</sup> Yang Liu,<sup>1,2,3</sup> Chi Wai Wong,<sup>8</sup> Tao Tan,<sup>9</sup> Weizhi Ji,<sup>9</sup> Patrick H. Maxwell<sup>12</sup>, Huanming Yang,<sup>1,10</sup> Jian Wang,<sup>1,10</sup> Shiping Liu,<sup>1,2,13</sup> Xun Xu,<sup>1,2,11,\*</sup> Yong Hou,<sup>1,2,13,\*</sup> Miguel A. Esteban,<sup>4,\*</sup> and Longqi Liu<sup>1,2,14,\*</sup>

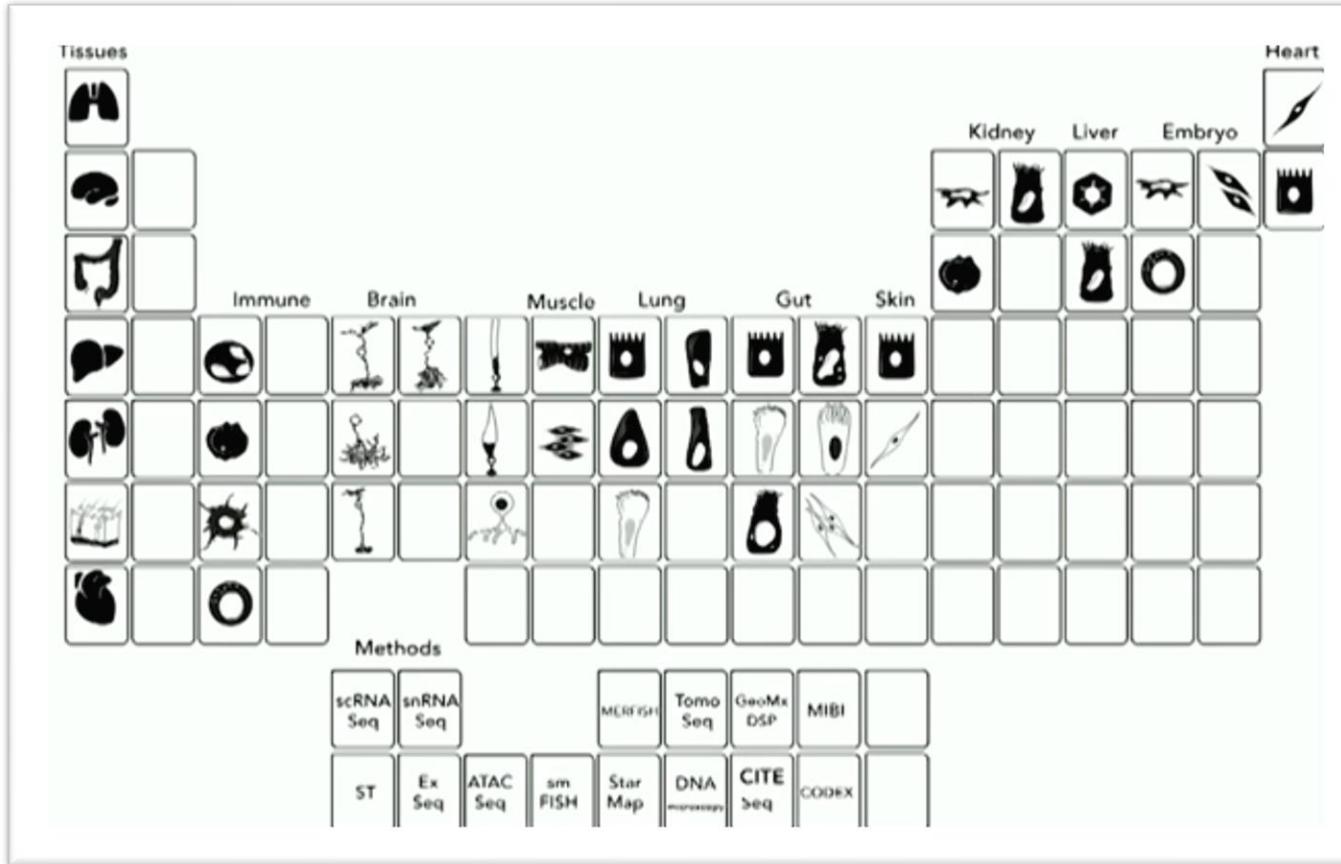


非人灵长类动物细胞图谱

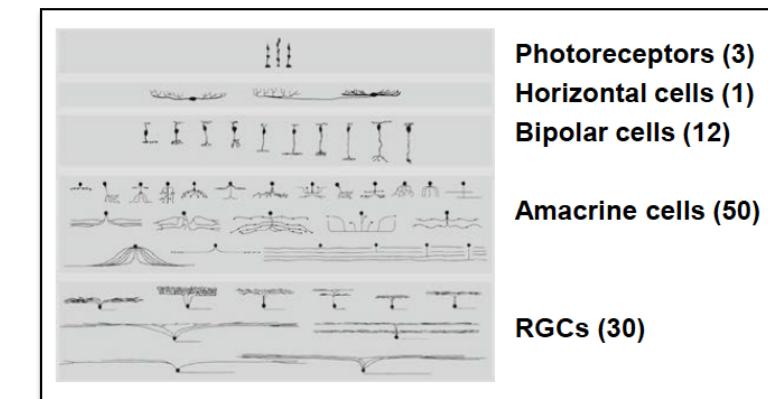
# 新冠病毒受体表达图谱



## “细胞周期表”



- Adult human:  $\sim 2 \times 10^{13}$  cells (excluding red blood cells)
- 300 'major' cell types
- .but,  $\sim 100$  sub-sub-types just of retinal neurons



Aviv, Sarah, et al., *The HCA whitepaper*, 2017

- 2013年, Nature Methods 杂志评选为生命科学领域年度技术
- 2018年, Science十大科学突破
- 2019年, 再次被Nature Methods杂志评选为2019年度技术
- 2020年, Nature: Technologies to watch in 2020, 单细胞测序为重点关注技术



## Method of the Year 2019: Single-cell multimodal omics

Multimodal omics measurement offers opportunities for gaining holistic views of cells one by one.

Editorial | Published: 30 December 2013

## Method of the Year 2013

Nature Methods 11, 1(2014) | Cite this article

3286 Accesses | 19 Citations | 109 Altmetric | Metrics

Methods to sequence the DNA and RNA of single cells are poised to transform many areas of biology and medicine.



# 2018

BREAKTHROUGH  
*of the YEAR*

## Development cell by cell

With a trio of techniques, scientists are tracking embryo development in stunning detail

TECHNOLOGY FEATURE · 21 JANUARY 2020

## Technologies to watch in 2020

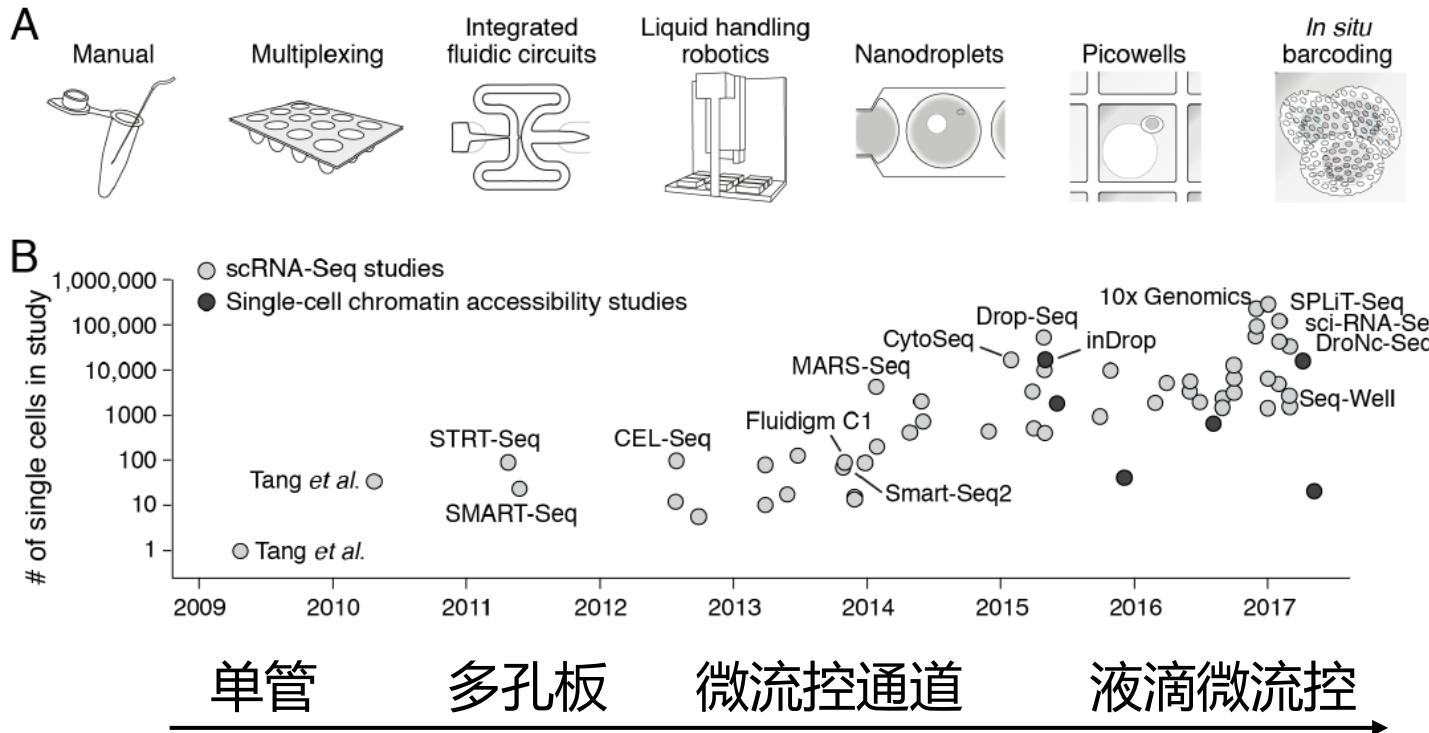
Thought leaders predict the tech developments that could have a big impact in the coming year.

### Single-cell sequencing

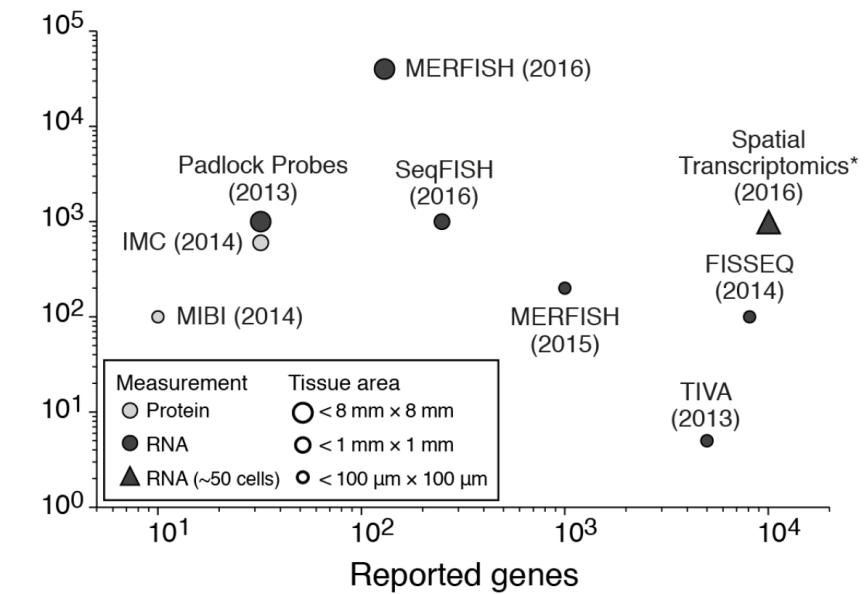
**J. Christopher Love** Chemical engineer at the Koch Institute for Integrative Cancer Research at MIT in Cambridge, Massachusetts.



## 单细胞测序技术的10年发展

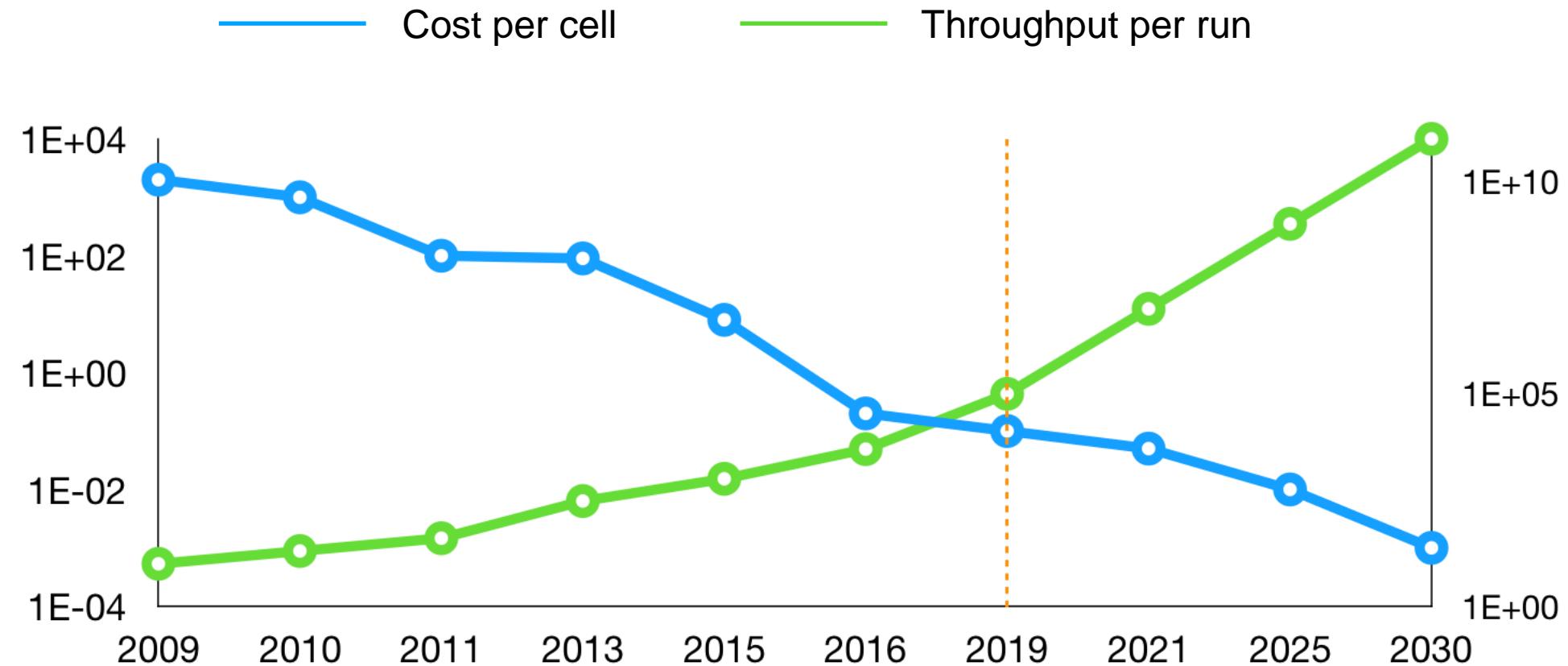


## 空间组学发展5年



HCA white book, 2017

过去10年来，细胞组学技术不断升级，灵敏度，通量都在不断上升，这些突破正带来发育，疾病，脑科学，生命演化等研究领域的全面变革。



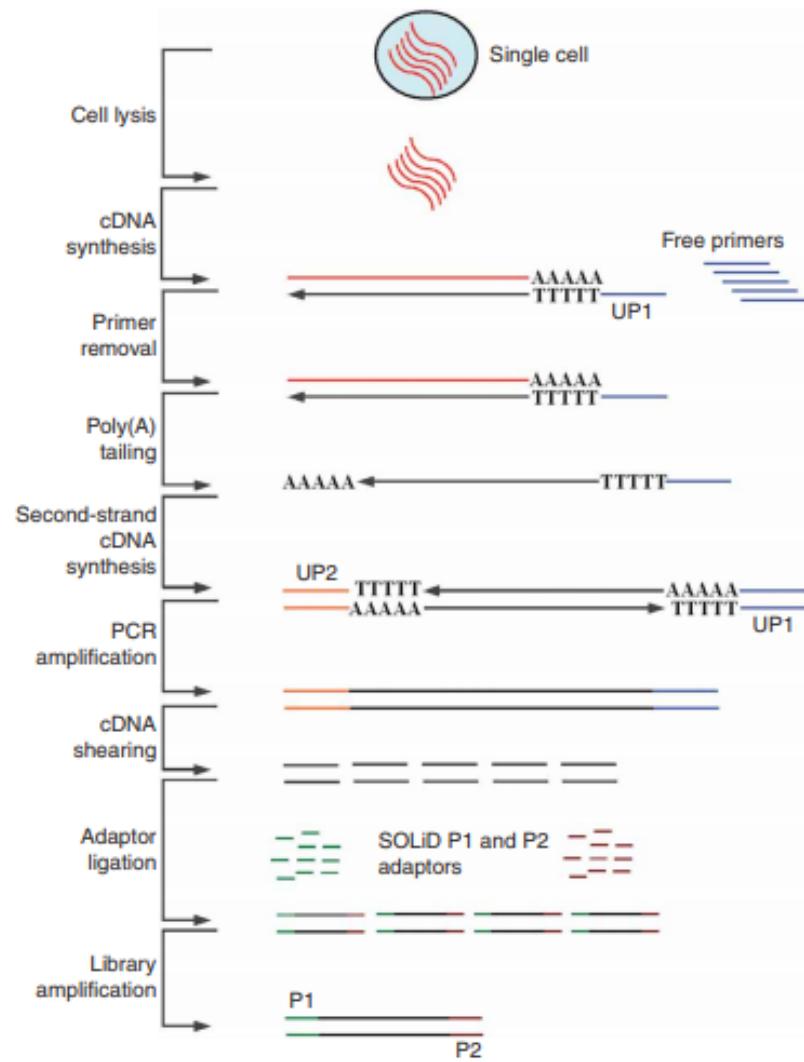
细胞组学技术通量不断上升，成本在以超摩尔定律下降，基本每两年会实现一个数量级的变化

Article | Published: 06 April 2009

## mRNA-Seq whole-transcriptome analysis of a single cell

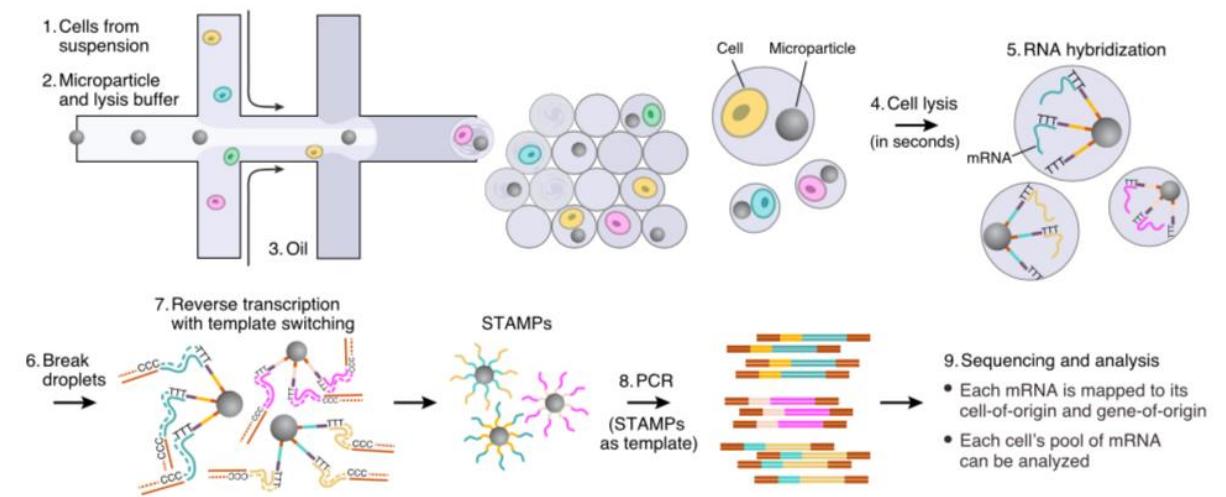
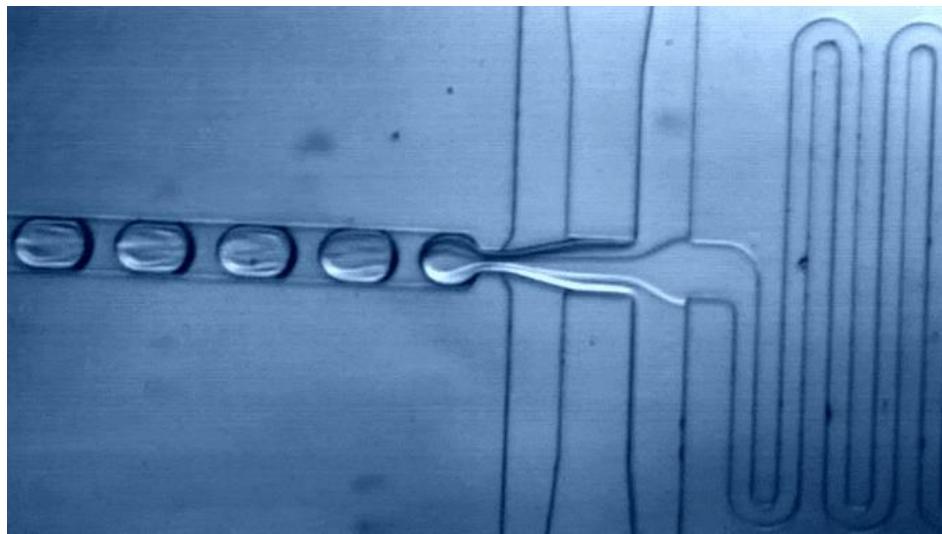
Fuchou Tang, Catalin Barbacioru, Yangzhou Wang, Ellen Nordman, Clarence Lee, Nanlan Xu, Xiaohui Wang, John Bodeau, Brian B Tuch, Asim Siddiqui, Kaiqin Lao & M Azim Surani

Next-generation sequencing technology is a powerful tool for transcriptome analysis. However, under certain conditions, only a small amount of material is available, which requires more sensitive techniques that can preferably be used at the single-cell level. Here we describe a single-cell digital gene expression profiling assay. Using our mRNA-Seq assay with only a single mouse blastomere, we detected the expression of 75% (5,270) more genes than microarray techniques and identified 1,753 previously unknown splice junctions called by at least 5 reads. Moreover, 8–19% of the genes with multiple known transcript isoforms expressed at least two isoforms in the same blastomere or oocyte, which unambiguously demonstrated the complexity of the transcript variants at whole-genome scale in individual cells. Finally, for *Dicer1<sup>-/-</sup>* and *Ago2<sup>-/-</sup>* (*Eif2c2<sup>-/-</sup>*) oocytes, we found that 1,696 and 1,553 genes, respectively, were abnormally upregulated compared to wild-type controls, with 619 genes in common.



# 高通量单细胞测序

BGI 华大

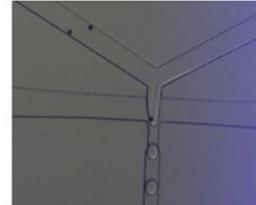


# Large-scale single-cell genomics center of BGI

Automated Sample Prep.



High-Throughput Single-Cell Barcoding



Automated Single-Cell Library Prep.



Large-Scale Single-Cell Sequencing

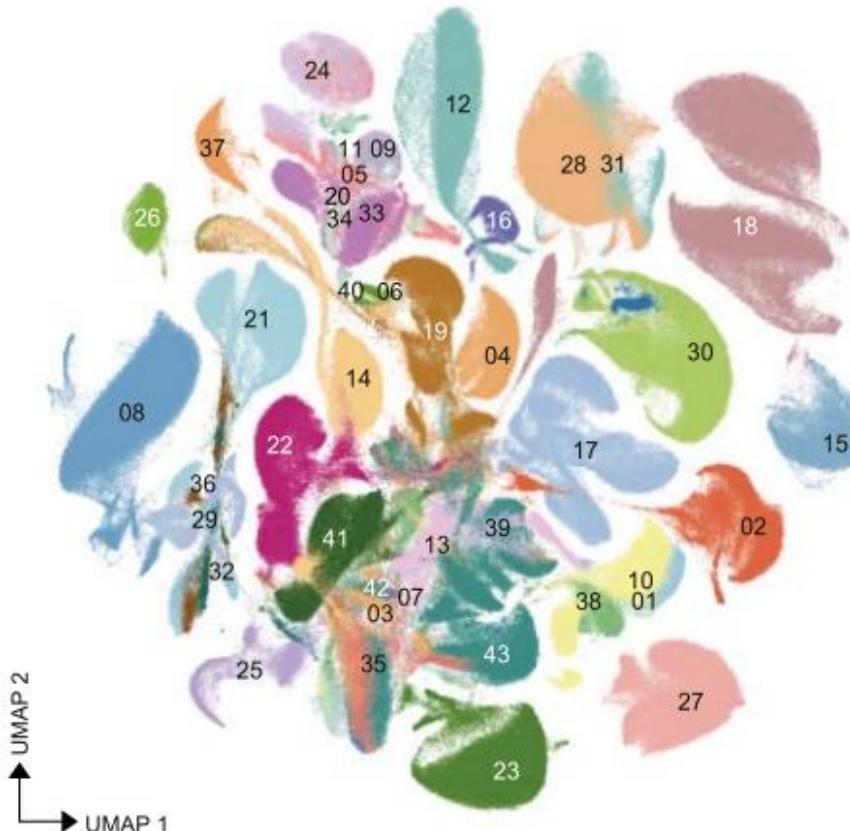
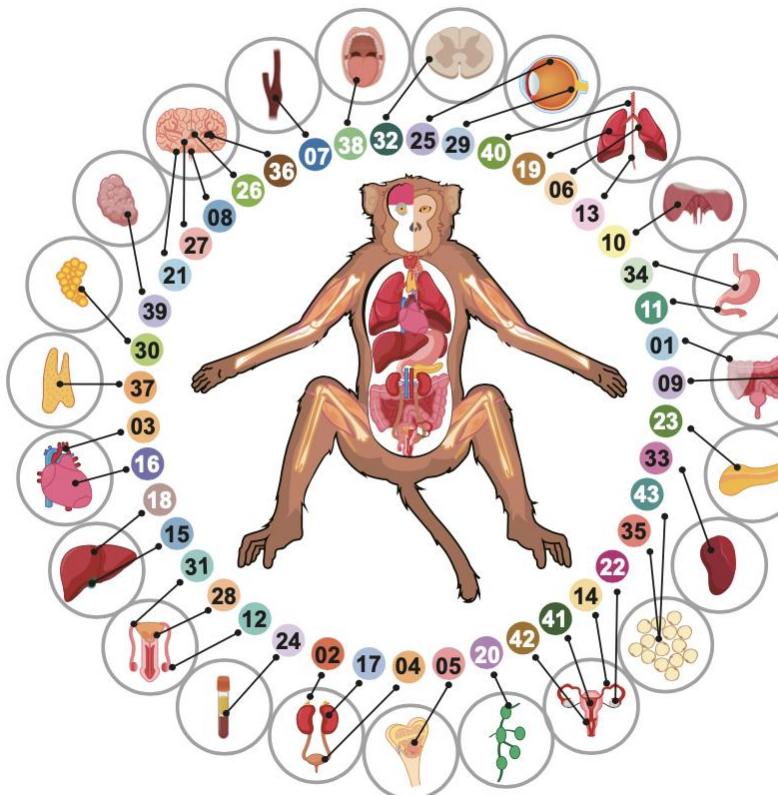


- ❖ 200,000 single-cells per day
- ❖ 50 million single-cell datasets per year
- ❖ 30 P sequencing data per year



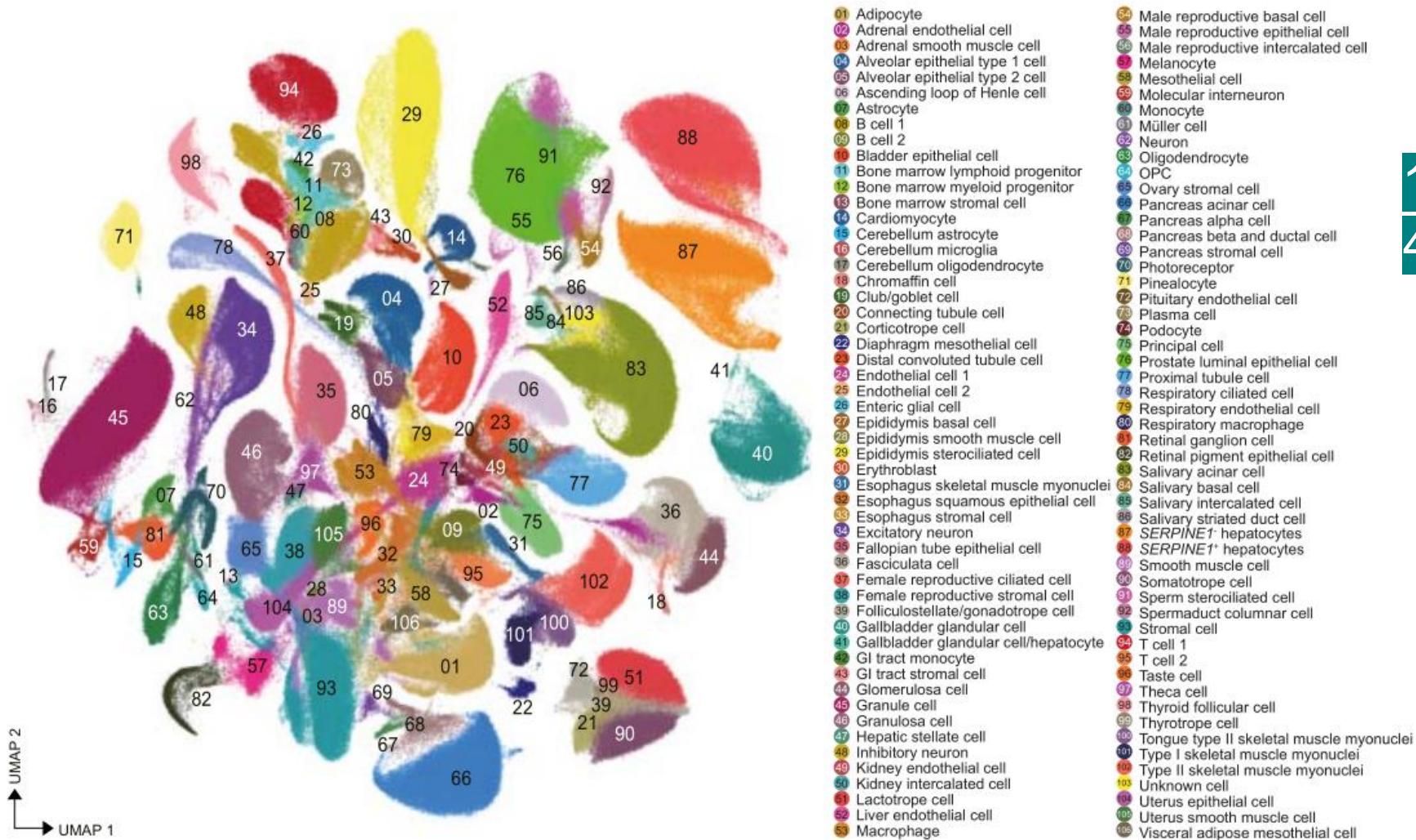
# Construction of a single-cell atlas across 43 tissues of a *Macaca fascicularis* monkey

Total cell number: 1,084,164



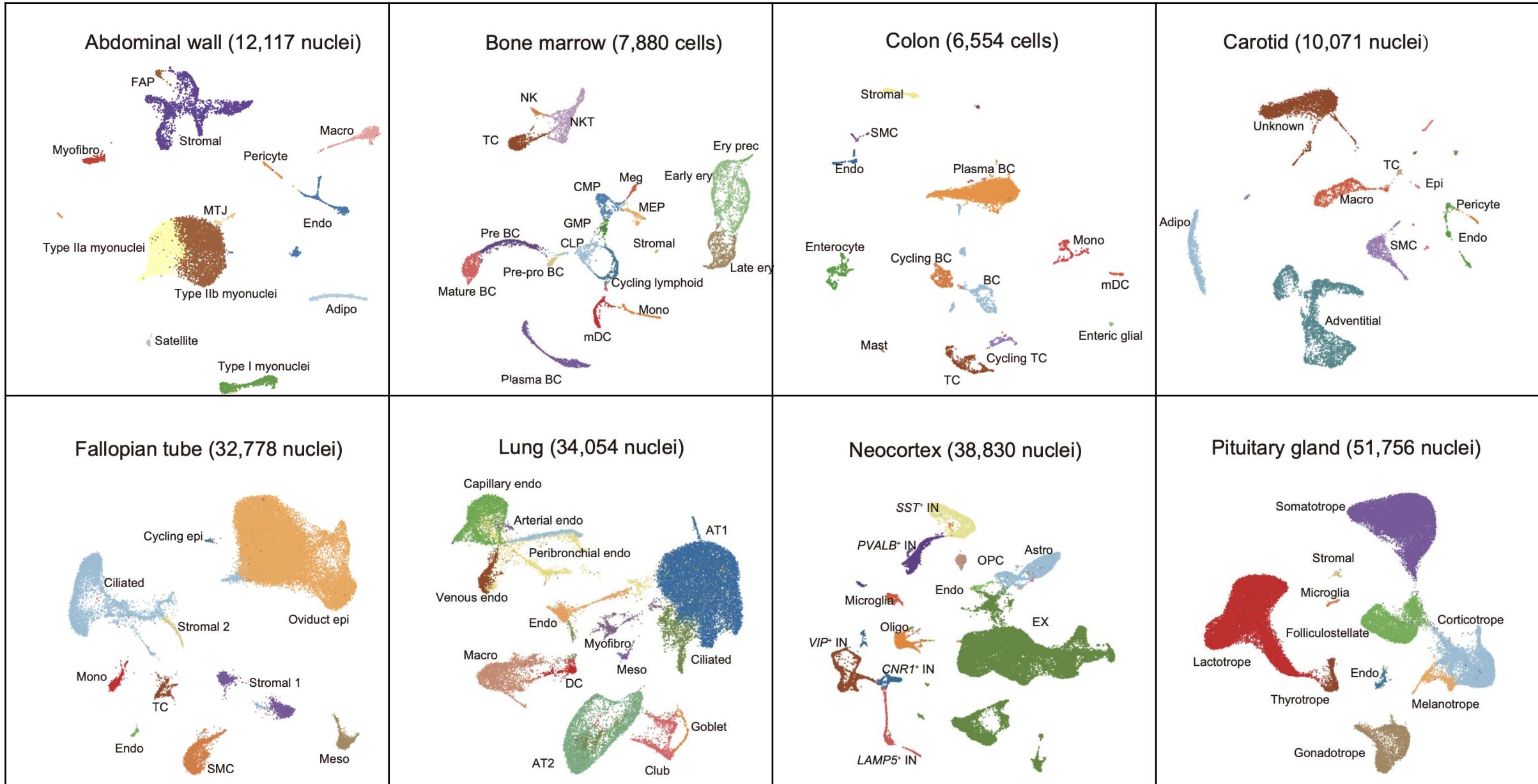
- |                   |               |                   |               |                   |                    |                         |                     |                     |
|-------------------|---------------|-------------------|---------------|-------------------|--------------------|-------------------------|---------------------|---------------------|
| 01 Abdominal wall | 06 Bronchus   | 11 Duodenum       | 16 Heart      | 21 Neocortex      | 26 Pineal gland    | 31 Spermaduct           | 36 Substantia nigra | 41 Ute              |
| 02 Adrenal gland  | 07 Carotid    | 12 Epididymis     | 17 Kidney     | 22 Ovary          | 27 Pituitary gland | 32 Spinal cord          | 37 Thyroid          | 42 Va_              |
| 03 Aorta          | 08 Cerebellum | 13 Esophagus      | 18 Liver      | 23 Pancreas       | 28 Prostate        | 33 Spleen               | 38 Tongue           | 43 Visceral adipose |
| 04 Bladder        | 09 Colon      | 14 Fallopian tube | 19 Lung       | 24 PBMC           | 29 Retina          | 34 Stomach              | 39 Tonsil           | 40 Trachea          |
| 05 Bone marrow    | 10 Diaphragm  | 15 Gallbladder    | 20 Lymph node | 25 Pigmentary epi | 30 Salivary gland  | 35 Subcutaneous adipose |                     |                     |

# Global annotation cell types

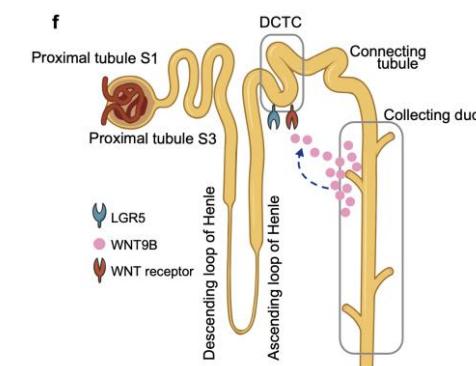
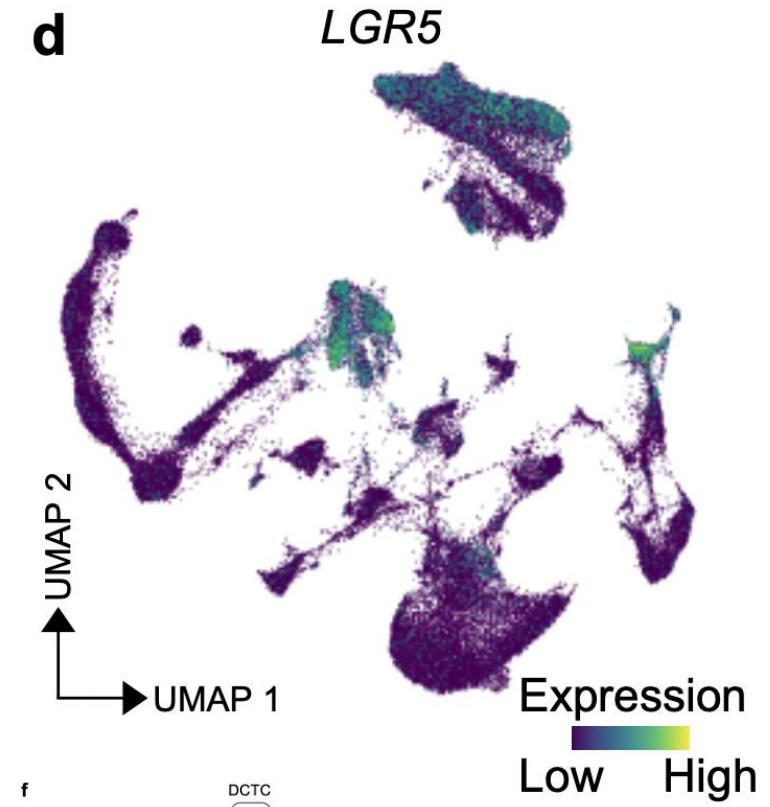
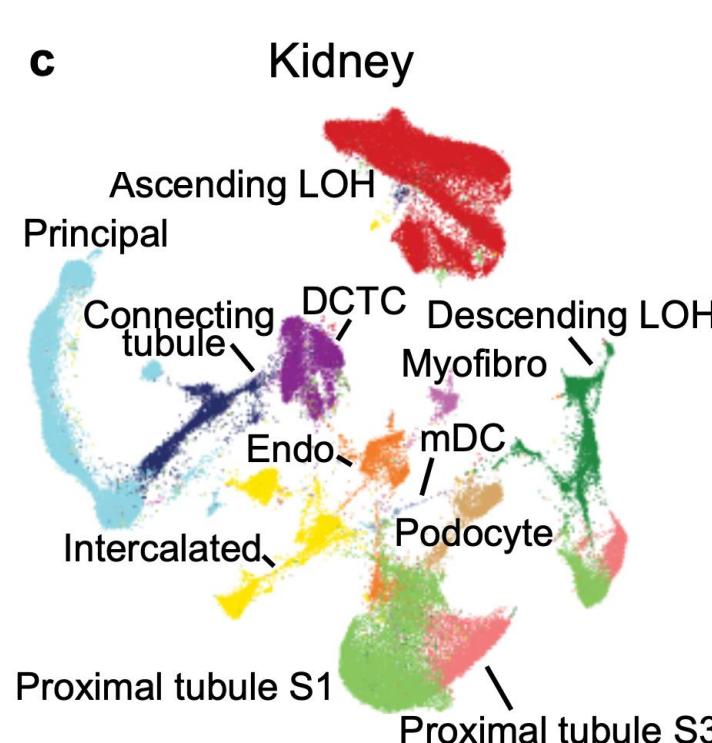
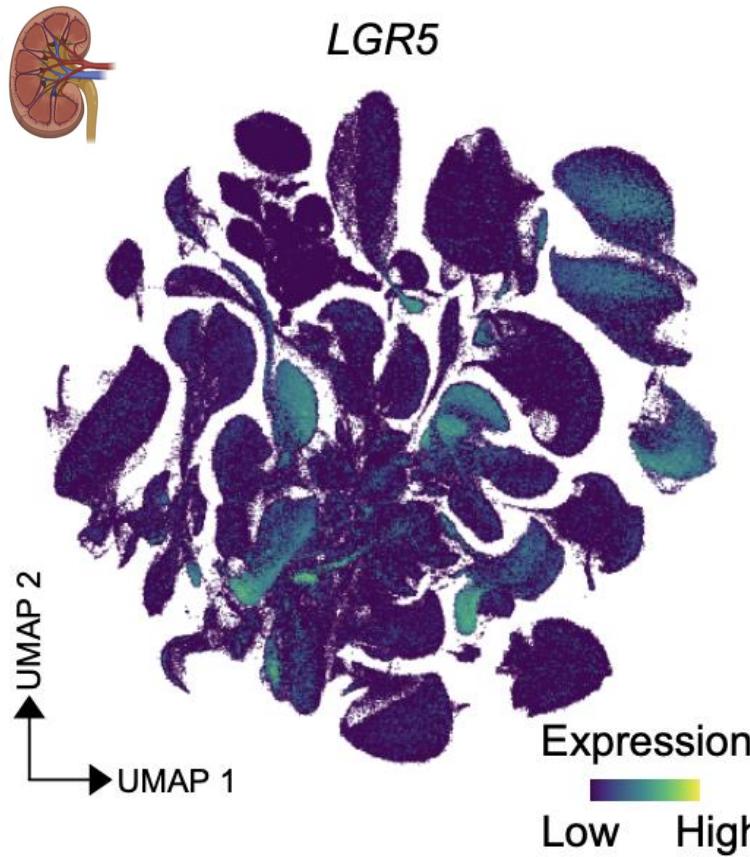


106 major cell types  
463 sub-cell types

# Clustering and annotation of representative



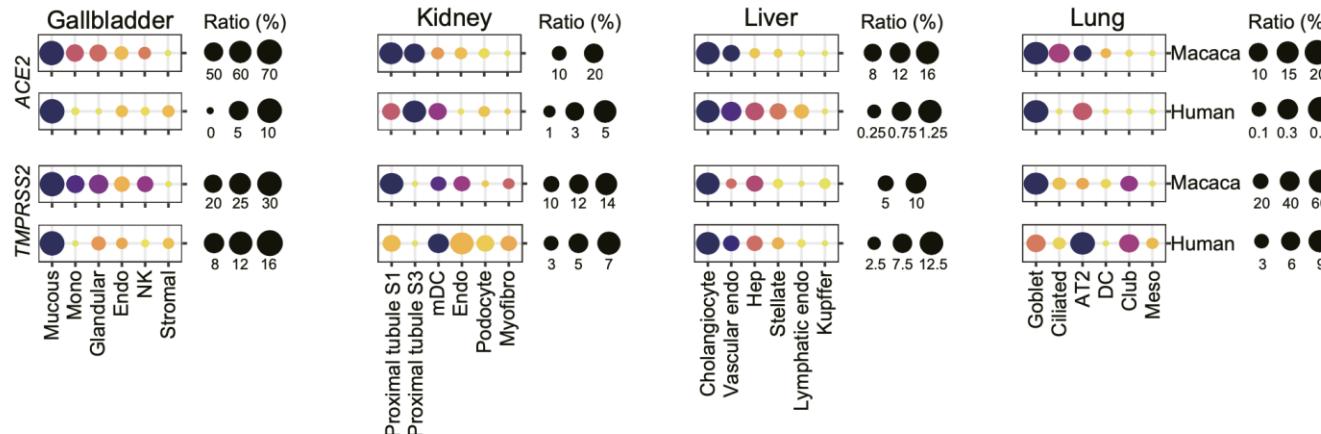
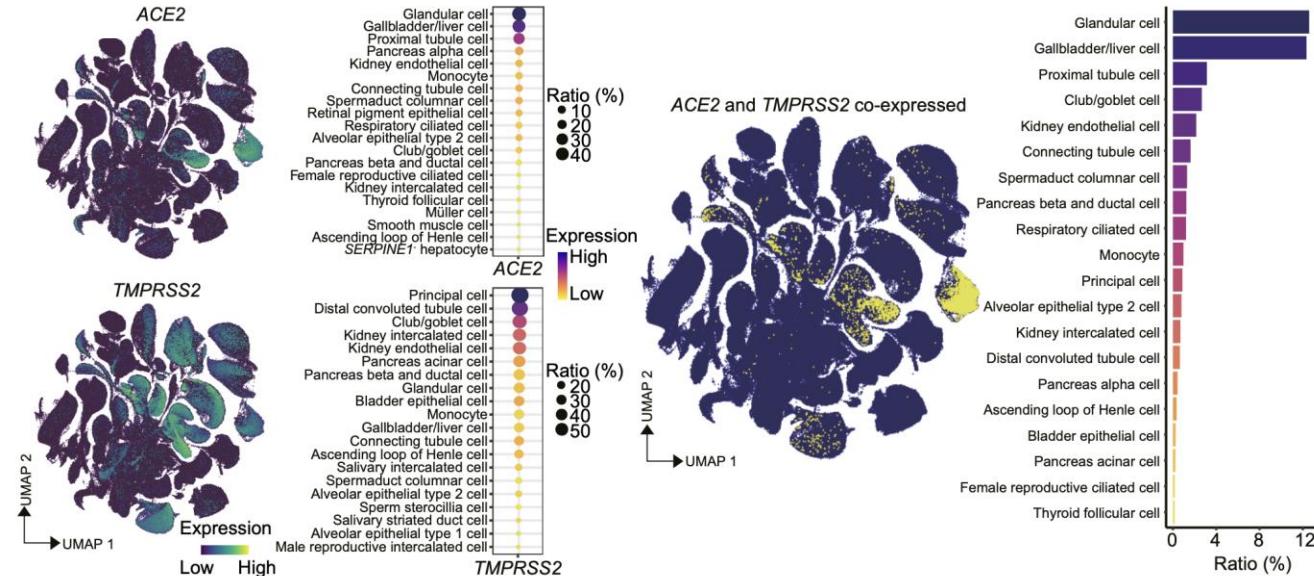
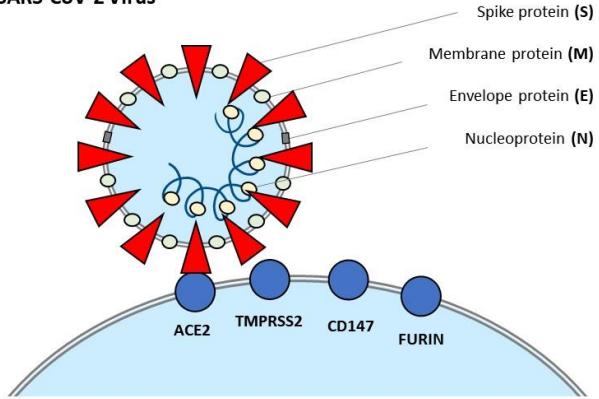
# Potential stem/progenitor cell in adult monkey tissues



Highest expression in type I skeletal muscle myonuclei,  
epithelial cells of the uterus and fallopian tube, oligodendrocyte  
progenitor cells (OPC) and renal distal convoluted tubule cells  
(DCTC)

# Profiling of SARS-CoV-2 receptor and co-receptor

The SARS-CoV-2 Virus



ARTICLE



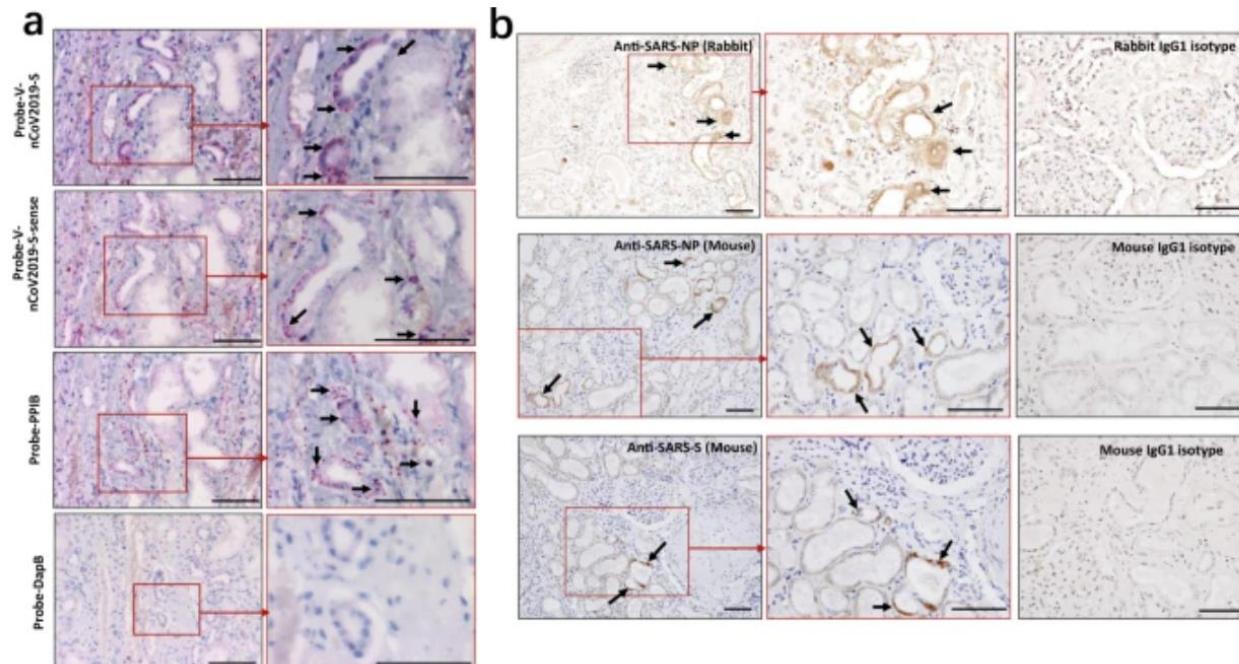
<https://doi.org/10.1038/s41467-021-22781-1>

OPEN

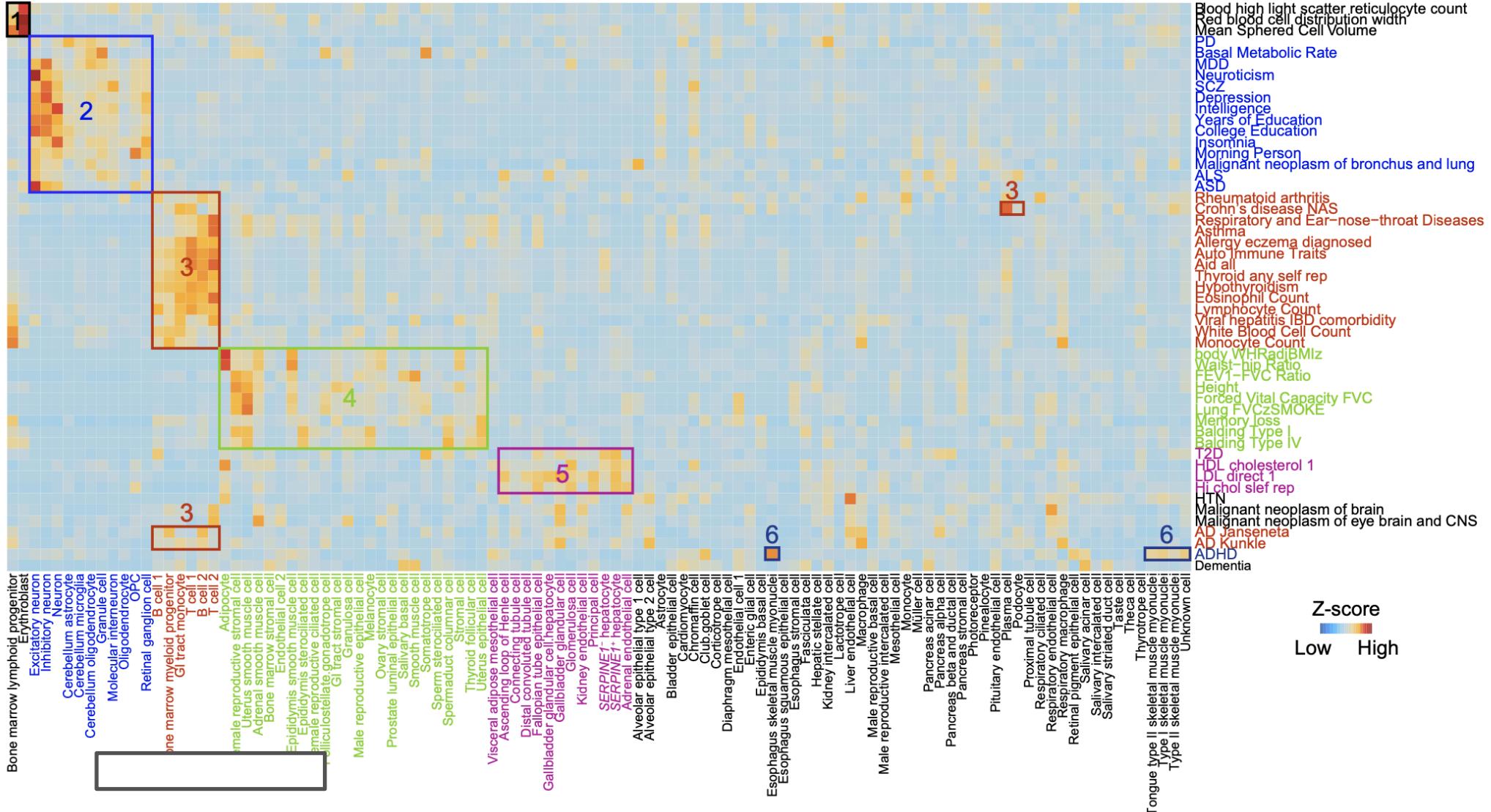
## Human kidney is a target for novel severe acute respiratory syndrome coronavirus 2 infection

Bo Diao<sup>1,2,3,9</sup>, Chenhui Wang<sup>1,9</sup>, Rongshuai Wang<sup>4,9</sup>, Zeqing Feng<sup>1</sup>, Ji Zhang<sup>1</sup>, Han Yang<sup>1</sup>, Yingjun Tan<sup>2</sup>, Huiming Wang<sup>5</sup>, Changsong Wang<sup>6</sup>, Liang Liu<sup>7</sup>, Ying Liu<sup>2</sup>, Yueping Liu<sup>2</sup>, Gang Wang<sup>2</sup>, Zilin Yuan<sup>2</sup>, Xiaotao Hou<sup>8</sup>, Liang Ren<sup>4,9</sup>✉, Yuzhang Wu<sup>1,9</sup>✉ & Yongwen Chen<sup>1,9</sup>✉

**Fig. 2: Detection of SARS-CoV-2 viral RNA and protein antigens in kidney tissues.**

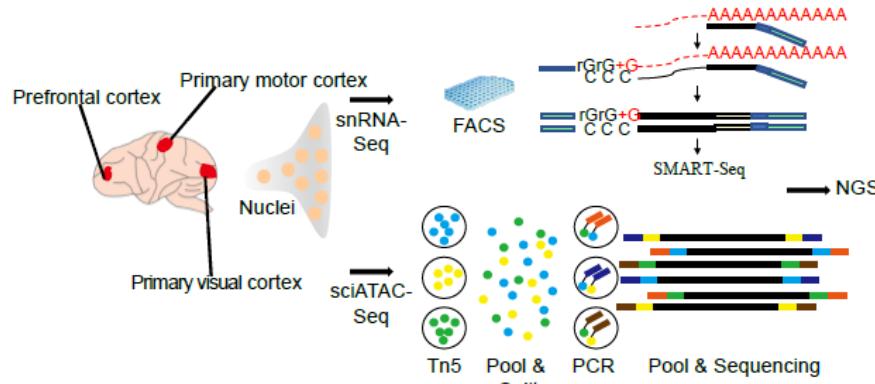
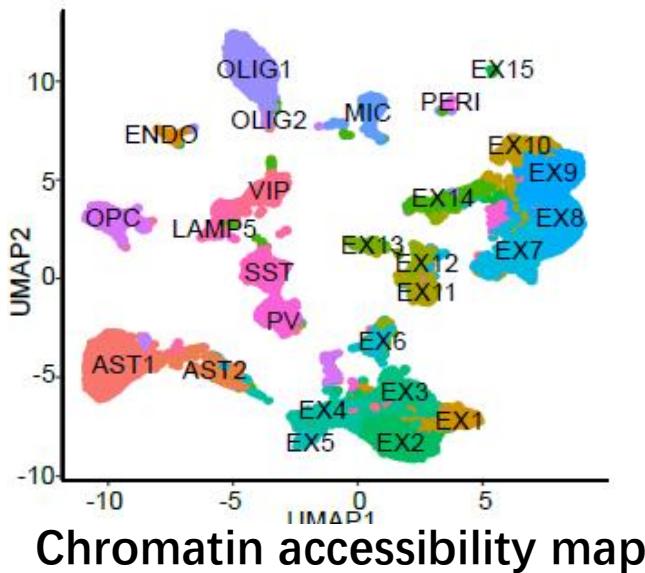


# Correlation of common human traits and genetic diseases with cell type in non-human primate

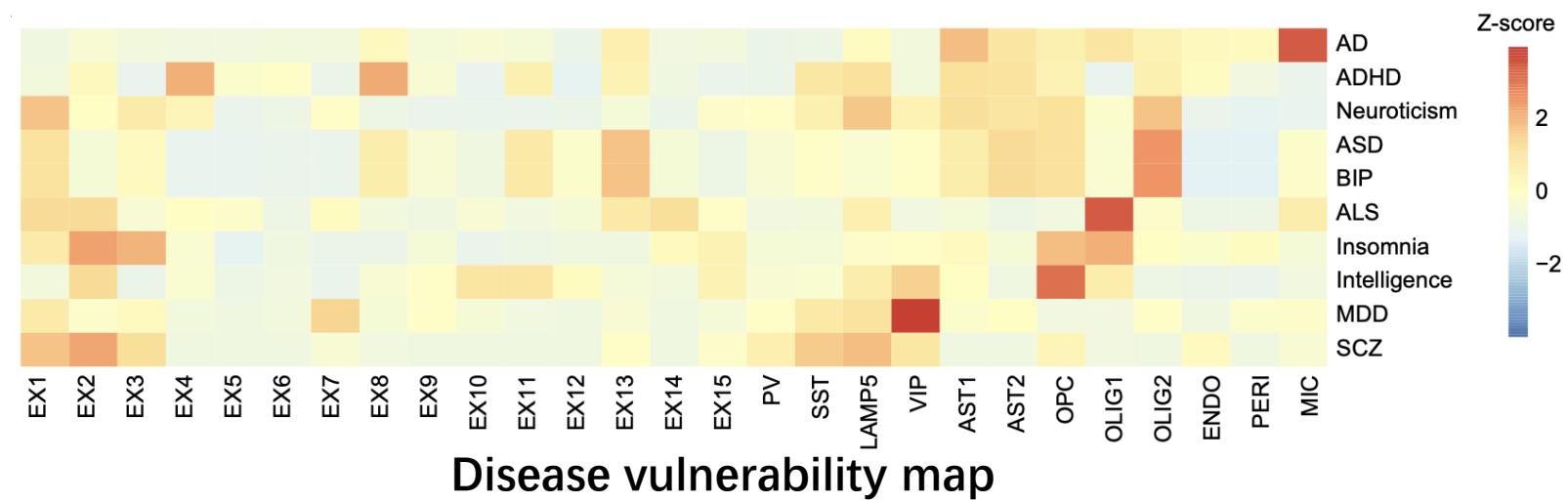


## Single-cell regulatory landscape and disease vulnerability map of adult Macaque cortex

Ying Lei<sup>1,13,18</sup>, Mengnan Cheng<sup>1,2,18</sup>, Zihao Li<sup>1,2,18</sup>, Zhenkun Zhuang<sup>1,3,18</sup>, Liang Wu<sup>1,2,18</sup>, Lei Han<sup>1,13</sup>, Shang Liu<sup>1,2</sup>, Zhihao Huang<sup>1</sup>, Jingkuan Wei<sup>4</sup>, Yuzhou Wang<sup>1</sup>, Zifei Wang<sup>1,2</sup>, Liqin Xu<sup>1</sup>, Taotao Pan<sup>1,2</sup>, Jiarui Xie<sup>1,3</sup>, Chuanyu Liu<sup>1</sup>, Giacomo Volpe<sup>5</sup>, Carl Ward<sup>5</sup>, Yiwei Lai<sup>5</sup>, Hao Yu<sup>1</sup>, Haixi Sun<sup>1</sup>, Qichao Yu<sup>1</sup>, Liang Wu<sup>1,14</sup>, Dandan Chen<sup>6</sup>, Chunqing Wang<sup>1,2</sup>, Chi Wai Wong<sup>7</sup>, Wei Liu<sup>7</sup>, Liangzhi Xu<sup>7</sup>, Zhouchun Shang<sup>1,8</sup>, Guibo Li<sup>1,14</sup>, Kun Ma<sup>1</sup>, Le Cheng<sup>9</sup>, Fei Ling<sup>3</sup>, Tao Tan<sup>4</sup>, Kai Chen<sup>4</sup>, Bosiljka Tasic<sup>10</sup>, Michael Dean<sup>11</sup>, Weizhi Ji<sup>4</sup>, Huanming Yang<sup>1,15</sup>, Miguel A. Esteban<sup>5,17</sup>, Xun Xu<sup>1,16</sup>, Hongkui Zeng<sup>10,\*</sup>, Longqi Liu<sup>1,17,\*</sup>, Yuyu Niu<sup>4,12,\*</sup>, Yong Hou<sup>1,14,\*</sup>, Shiping Liu<sup>1,14,19,\*</sup>



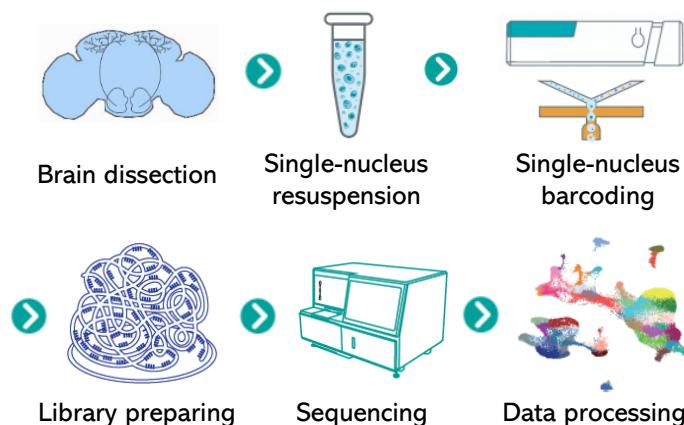
RNA: 11,477	ATAC: 43,793
PFC: 6,949	PFC: 14,480
M1: 1,971	M1: 9,853
V1: 2,557	V1: 19,460



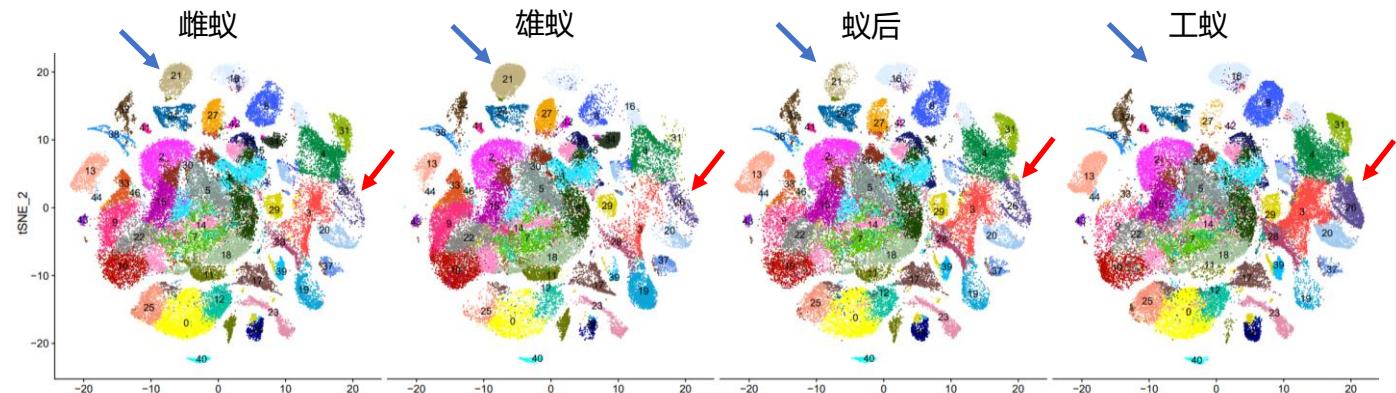
对食蟹猴初级视皮层、初级运动皮层以及前额皮质等3个区域脑组织中单细胞染色质可及性及基因表达进行了联合分析，分别鉴定出神经细胞和非神经细胞的转录组细胞亚型及表观组学细胞亚型

## 生物演化：解析蚂蚁社会分工的神经调控机制

## 解析昆虫的社会性演化



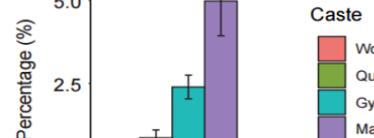
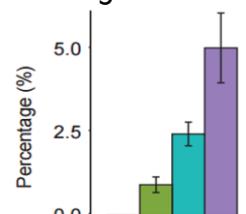
(华大研究院、昆明动物研究所等合作完成)



0	Unknown	7	Unknown	13	Kenyon cells	20	Unknown	27	Unknown	34	Monoaminergic neurons
1	Unknown	8	Kenyon cells	14	Kenyon cells	21	Kenyon cells	28	Unknown	35	Kenyon cells
2	Unknown	9	Unknown	15	Unknown	22	Ensheathing glia	29	Surface glia	36	Ensheathing glia
3	Unknown	10	Unknown	16	Unknown	23	Kenyon cells	30	Unknown	37	Kenyon cells
4	Unknown	11	Olfactory projection neurons	17	Astrocyte-like	24	Perineurial glia	31	Kenyon cells	38	Monoaminergic neurons
5	Kenyon cells	12	Unknown	18	Kenyon cells	25	Unknown	32	Perineurial glia	39	Cortex glia
6	Unknown	13	Kenyon cells	19	Unknown	26	Kenyon cells	33	Kenyon cells	40	Astrocyte-like

GO_ID	GO_Term	GO_Class
GO:0007617	mating behavior	BP
GO:0008049	male courtship behavior	BP
GO:0019098	reproductive behavior	BP

C16: Mating behavior-related neuron



在工蚁的大脑中，负责学习和记忆的KC细胞数量显著高于其他三个阶级，这种KC细胞的高度富集很可能就是工蚁具有多样的劳作行为的原因。一类未知类型的神经元，在工蚁大脑中完全缺失了，结果显示这类细胞的高表达基因功能主要负责生殖行为的调控。结合工蚁的阶级分工，也就是完全没有生殖行为，这类细胞很可能是工蚁没有生殖行为的原因。

- 
- 01** 基因组与细胞命运调控
  - 02** 单细胞组学发展及应用
  - 03** 时空组学发展及应用

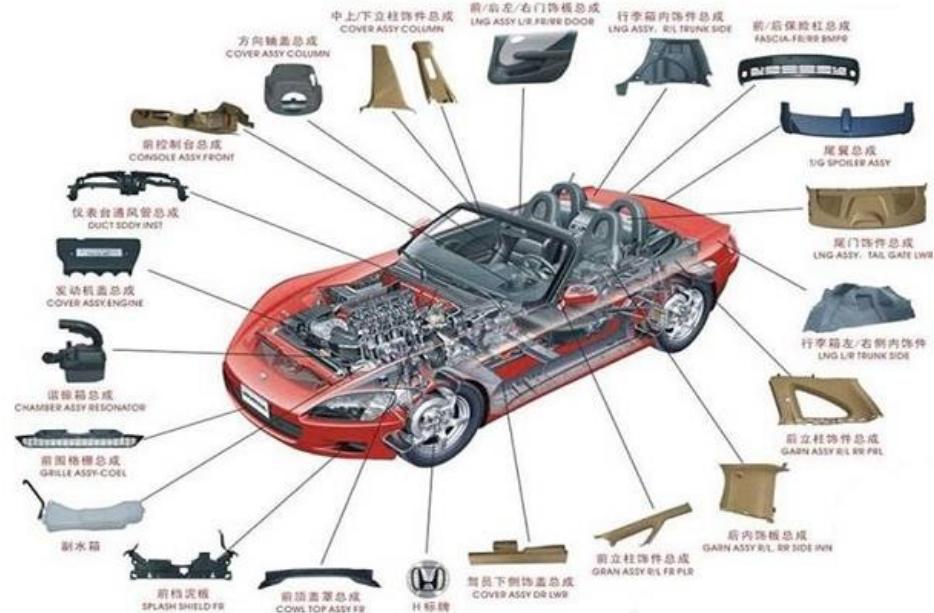


## | 单细胞组学 VS 空间组学

“单细胞组学”

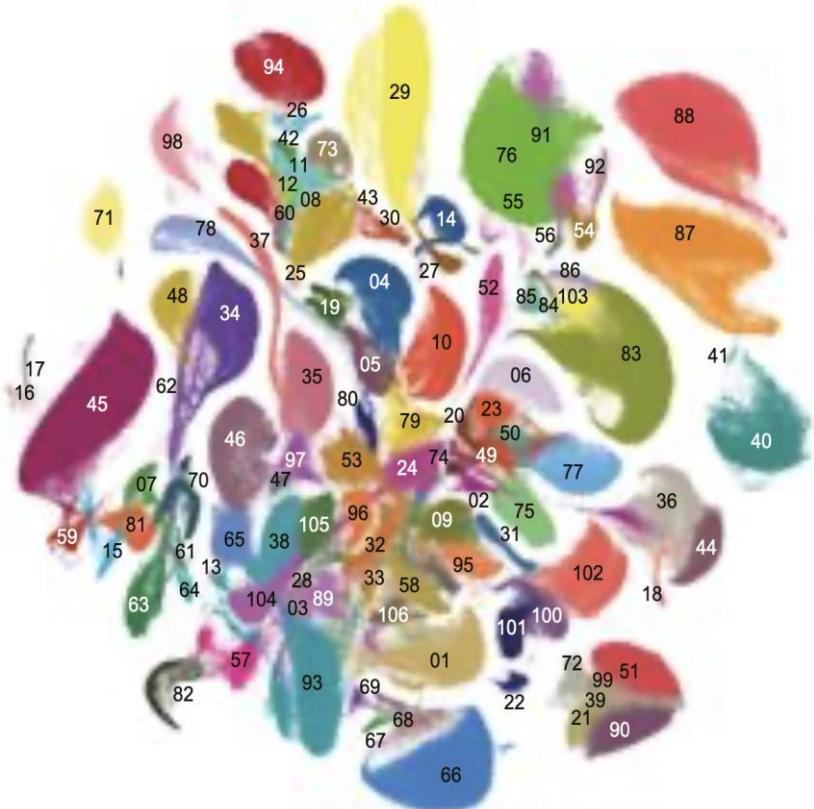


“空间组学”

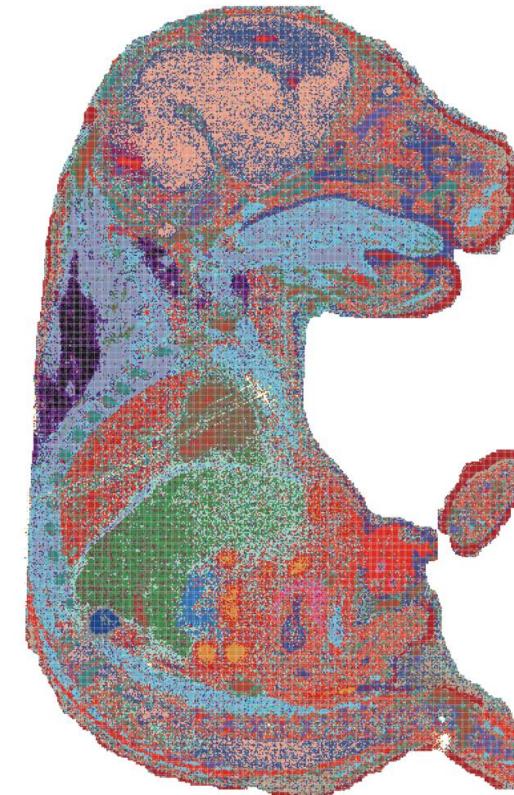


# | 单细胞组学 VS 空间组学

# 单细胞组学

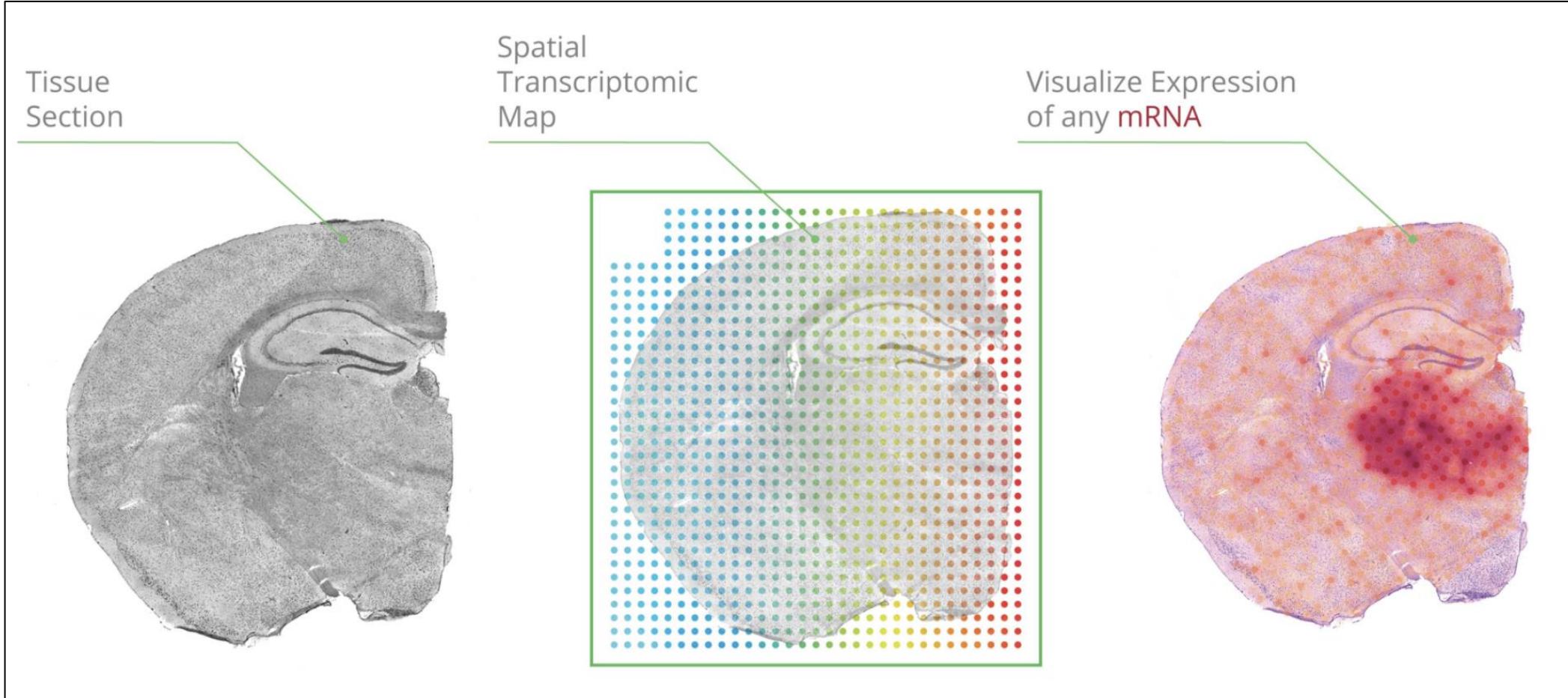


空间组学





- 在时间的尺度上，对生命体（或组织/器官）的不同组学信息（包括DNA, RNA, 蛋白质等单个组学或多组学）进行全基因组范围的、分子精度的空间定位。
- 时空组学可以用于分析细胞类型的空间定位，细胞的微环境，基因表达的空间定位，以及以上信息在时间尺度上的变化过程。





# 时空组学将带来生命科学领域第三次科技革命

显微镜



1665年

首次观察到生命微观结构-细胞

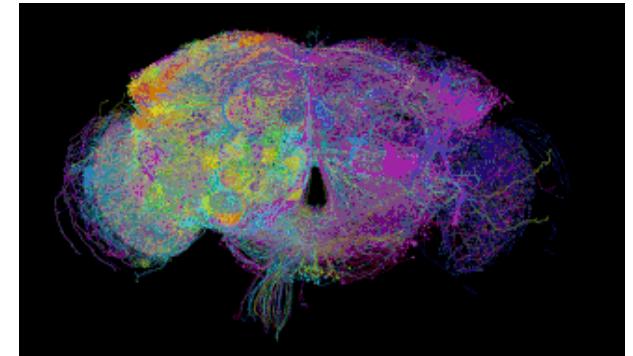
DNA测序技术



2001年

首次完成人类基因组图谱绘制

时空组学技术



2021年

首次实现生命全景时空分子图谱

## 跨尺度全景4D生命交响曲 4D Multi-dimensional Life Panorama



纳尺度 微观大场景 “因”  
Nanoscale Microcosm Cause

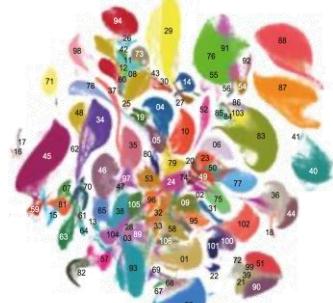
因果贯穿 跨尺度 全景 智能化  
Causality Multi-dimensional Panoramic Intelligent

“果” 微尺度 宏观大场景  
Outcome Microscale Macrography

A 基因 Gene  
B 表观 Appearance  
C 转录 Transcription  
D 蛋白质 Protein  
E 代谢组 Metabolome  
F 激素 Hormone  
G 免疫组化 Immunohistochemistry  
H 细胞 Cell

量子 · 原子 · 分子 Quantum · Atom · Molecule

J 亚细胞 Subcellular  
K 生化 Biochemistry  
L 生理 Physiology  
M 病理 Pathology  
N 组织 Tissue  
O 医学影像 Medical Imaging  
P 器官 Organ  
Q 个体 Individual



人类时空组  
跨尺度、全景式图谱

基因组学

细胞组学

解剖学

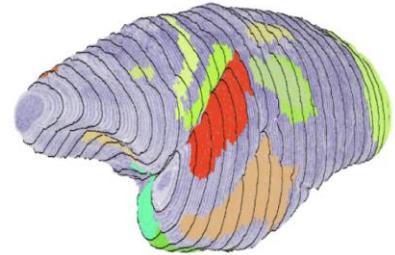
影像学



### 重新认识生命结构

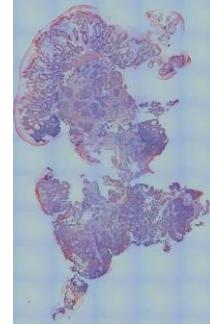


CT影像

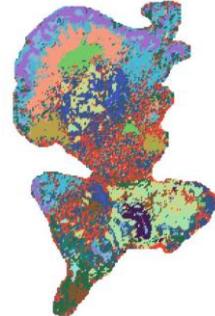


器官3D重构

### 重新定义疾病



免疫组化

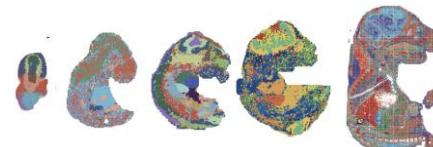


疾病3D数字化病理

### 重新认知个体发育

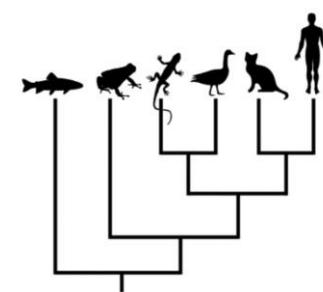


组织胚胎学染色

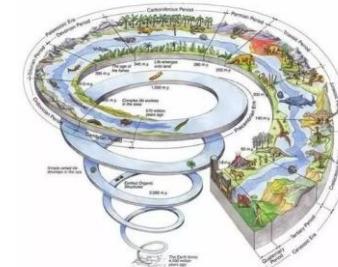


时空发育4D图谱

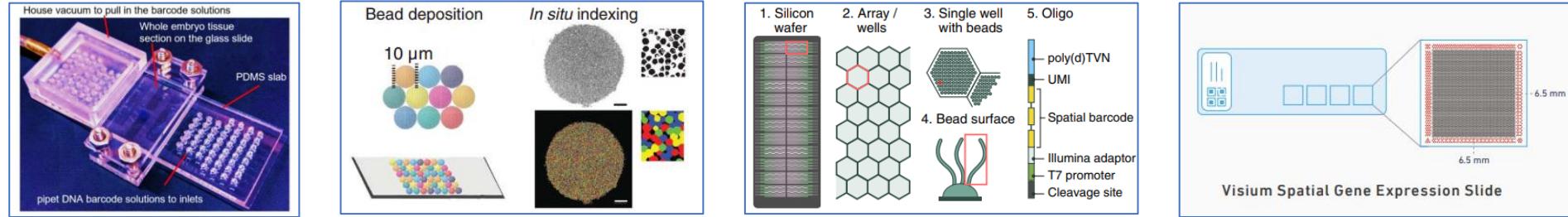
### 重新认知物种起源



物种进化树

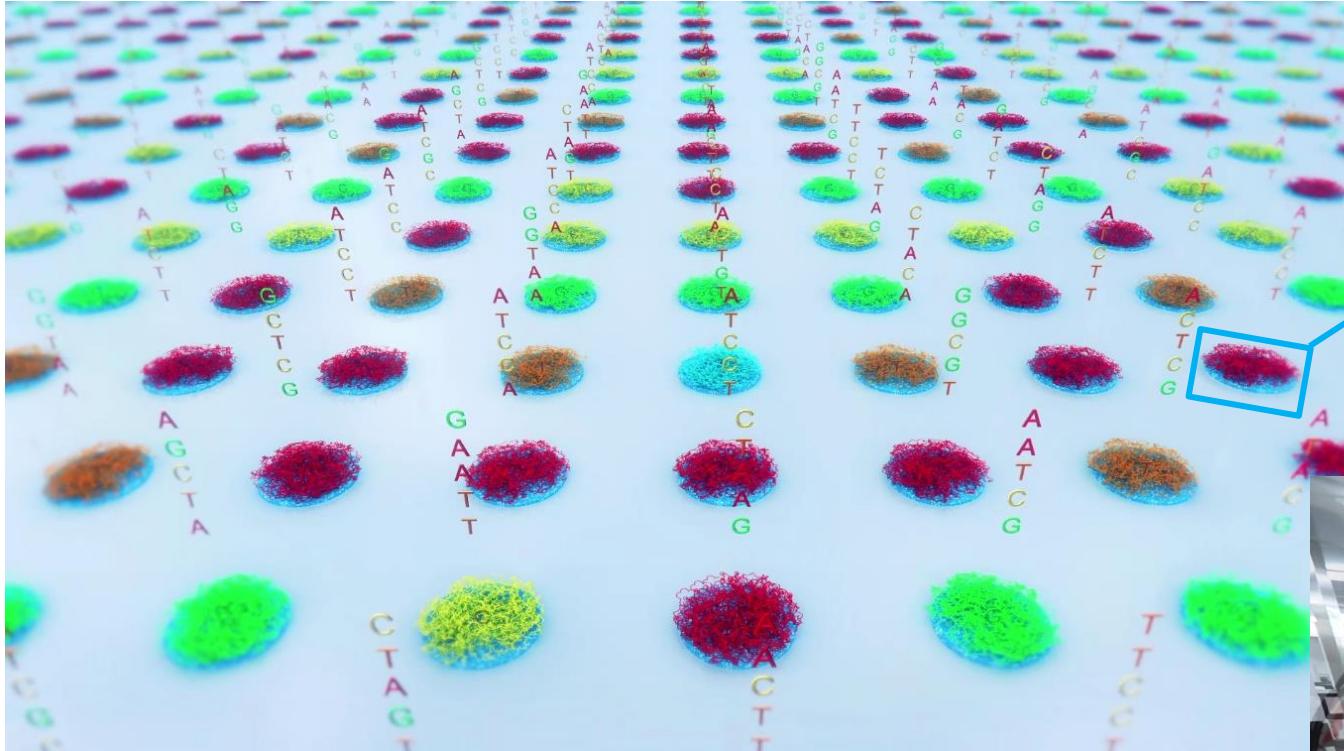


物种发育进化5D图谱



	DBiT-seq	Slide-seq	HDST	10x Genomics
信号点大小 (μm)	10	10	2	55
精度(μm)	20	10	2	100
视场 (mm)	1.0 x 1.0	Φ 3.0	5.7 x 2.4	6.5 x 6.5

精度和视场是当前技术的主要问题



### 纳米球印刷术

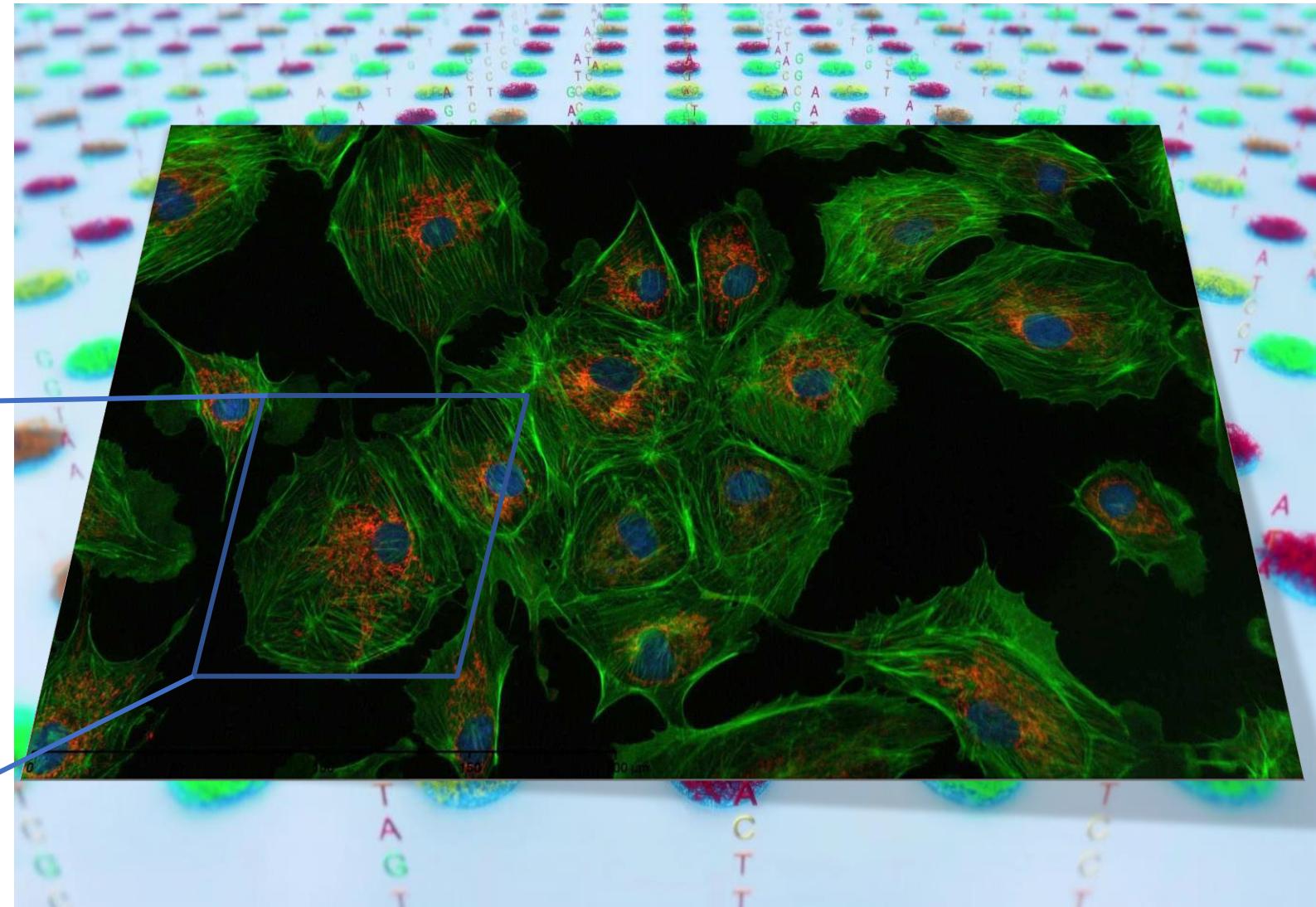
利用DNA纳米球坐标和序列信息进行“生化印刷”。

1. 坐标: FC ID, FOV ID, DNB ID
2. 序列: ATCCT



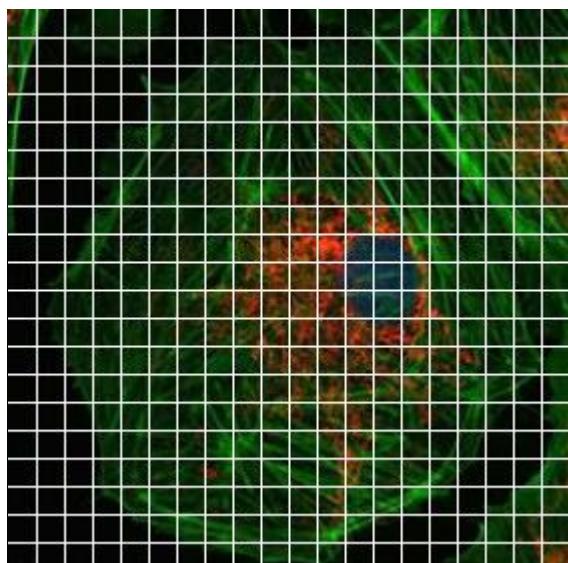
### 全景式测序

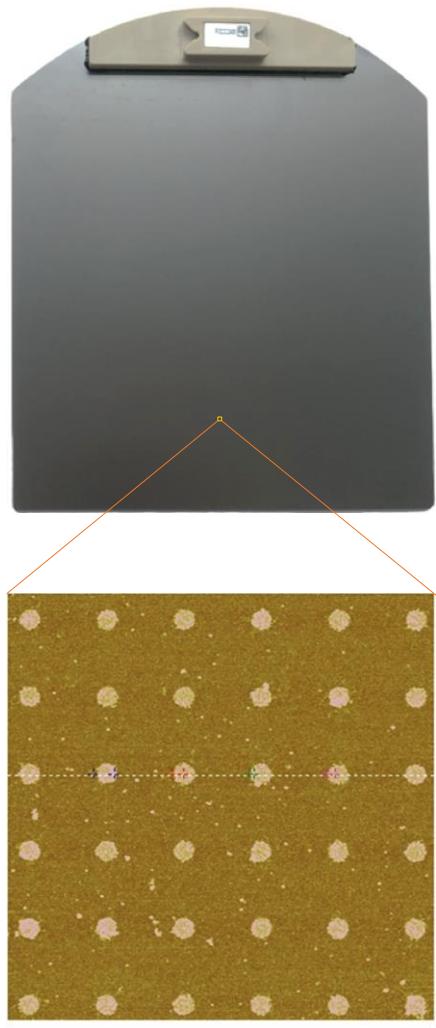
视场大小可达 **13cm \* 13 cm**



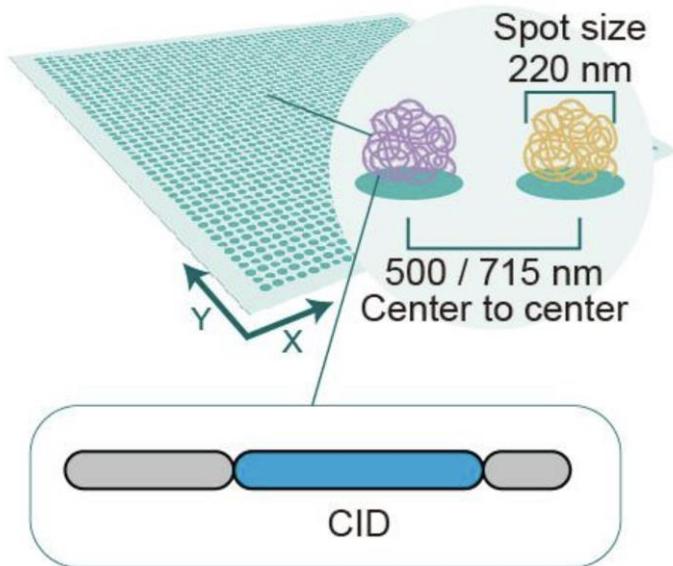
### 亚细胞定位

分辨率可达 **500nm**

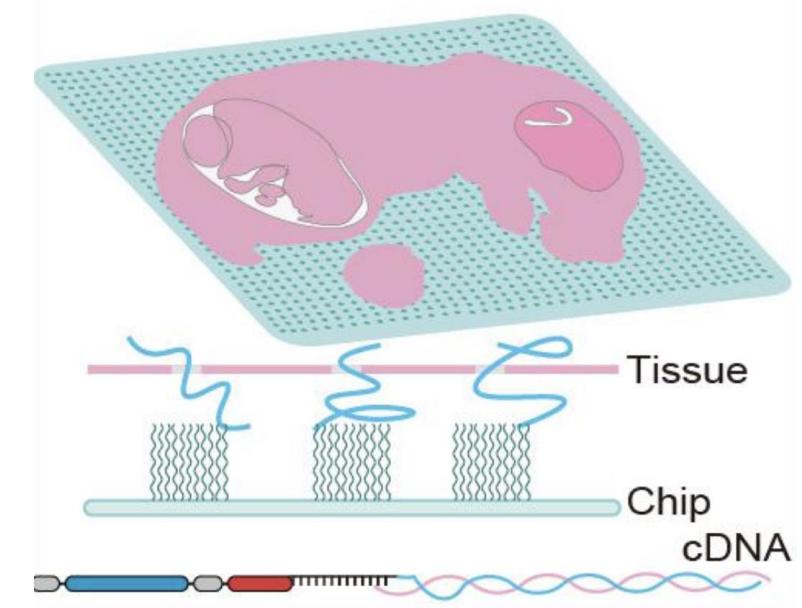




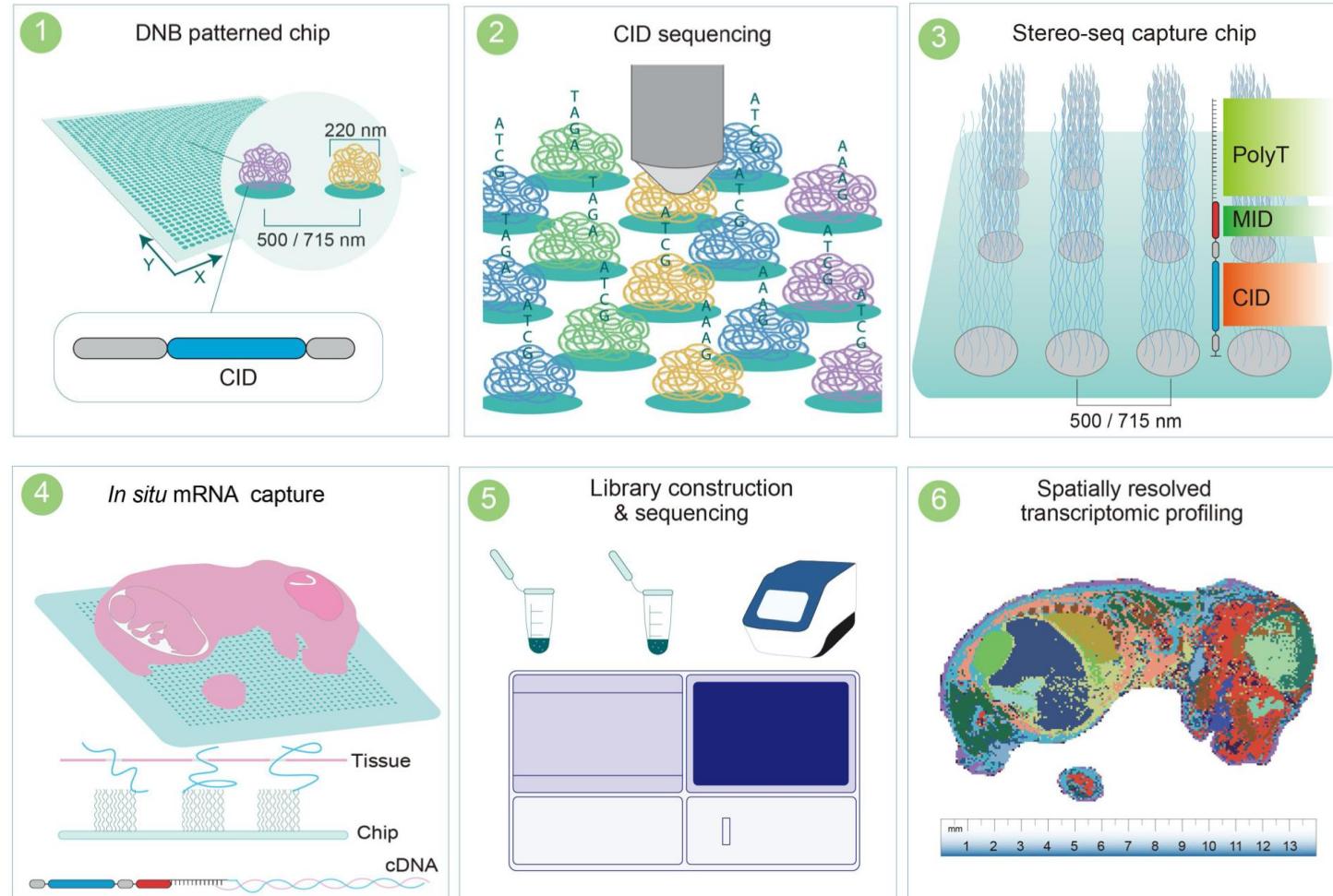
DNB纳米球芯片



组织原位捕获测序

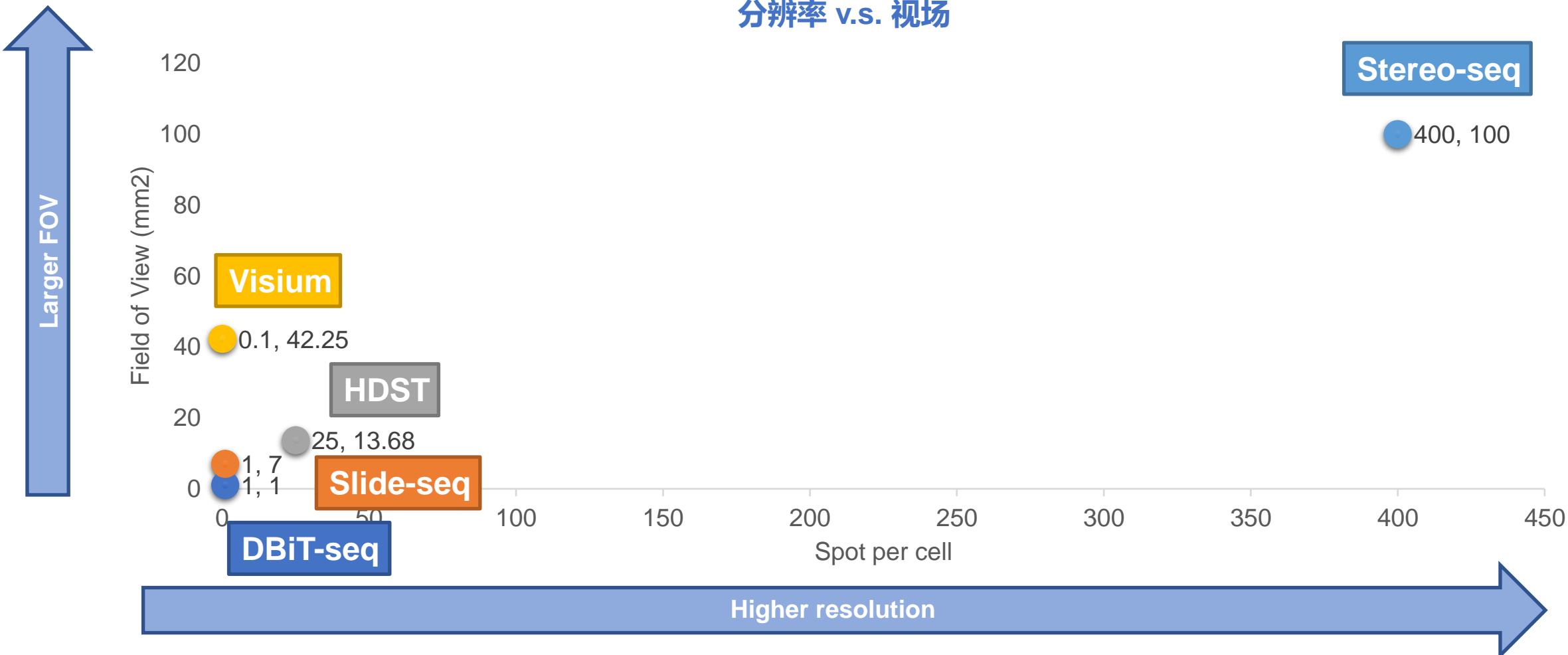


## 华大Stereo-seq时空技术流程

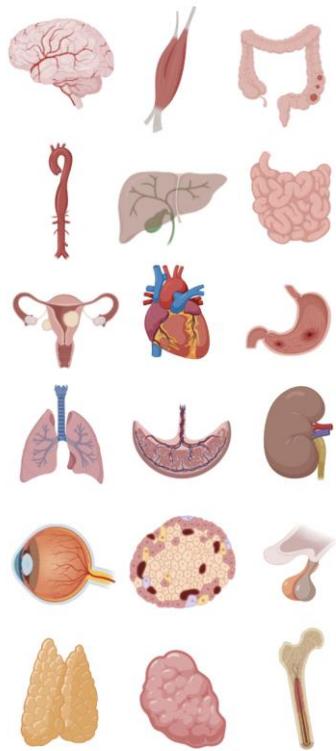


## 不同技术比较

分辨率 v.s. 视场



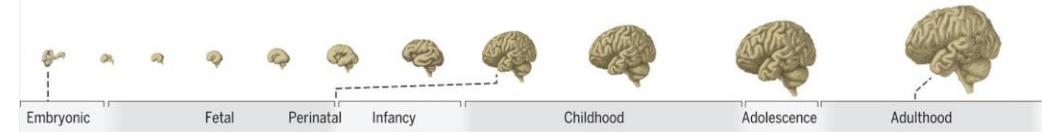
### 器官图谱



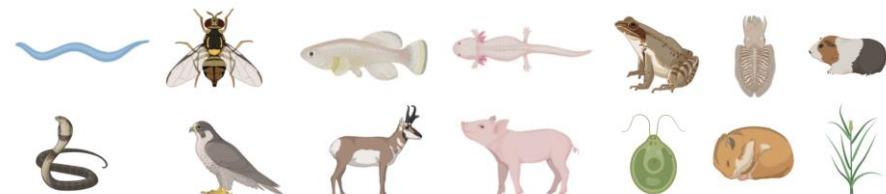
### 疾病图谱



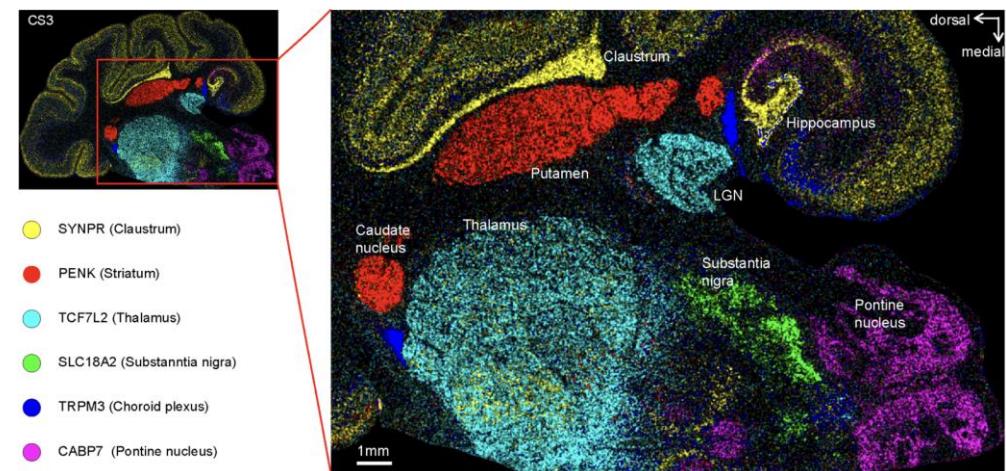
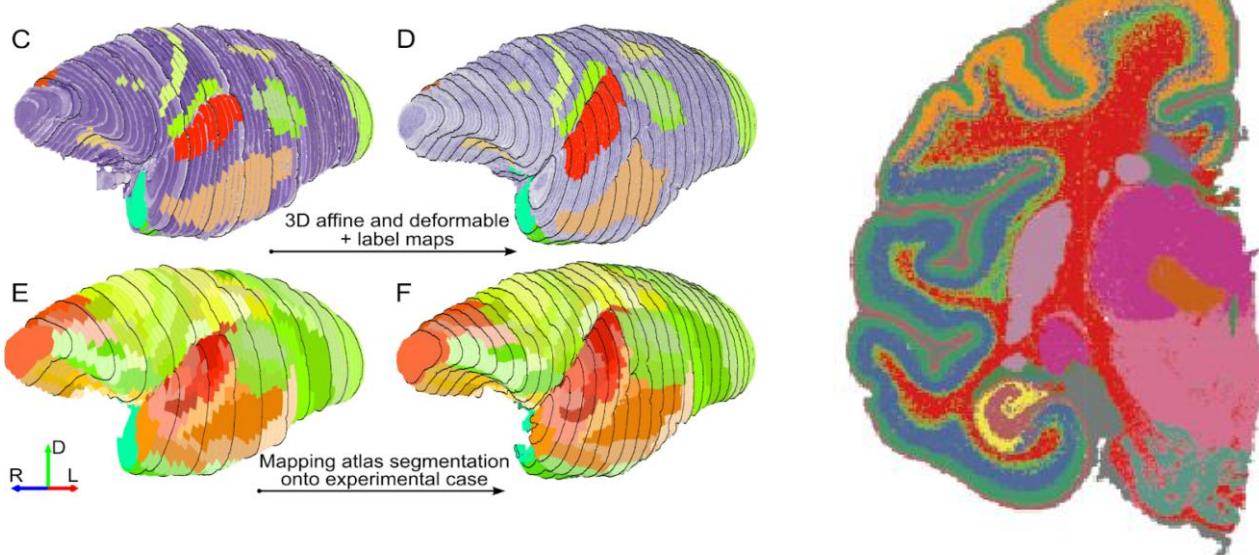
### 发育图谱



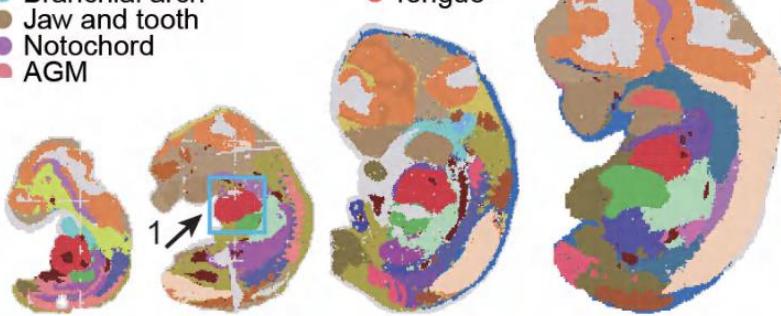
### 进化图谱



## 猴脑器官3D图谱



- Brain
- Spinal cord
- Neural crest
- Ganglion
- Facial ganglion
- Trigeminal ganglion
- Glossopharyngeal nerve
- Dorsal root ganglion
- Epidermis
- Hair follicle
- Mucosal epithelium
- Olfactory epithelium
- Sensory epithelium
- Head mesenchyme
- Meninges
- Branchial arch
- Jaw and tooth
- Notochord
- AGM

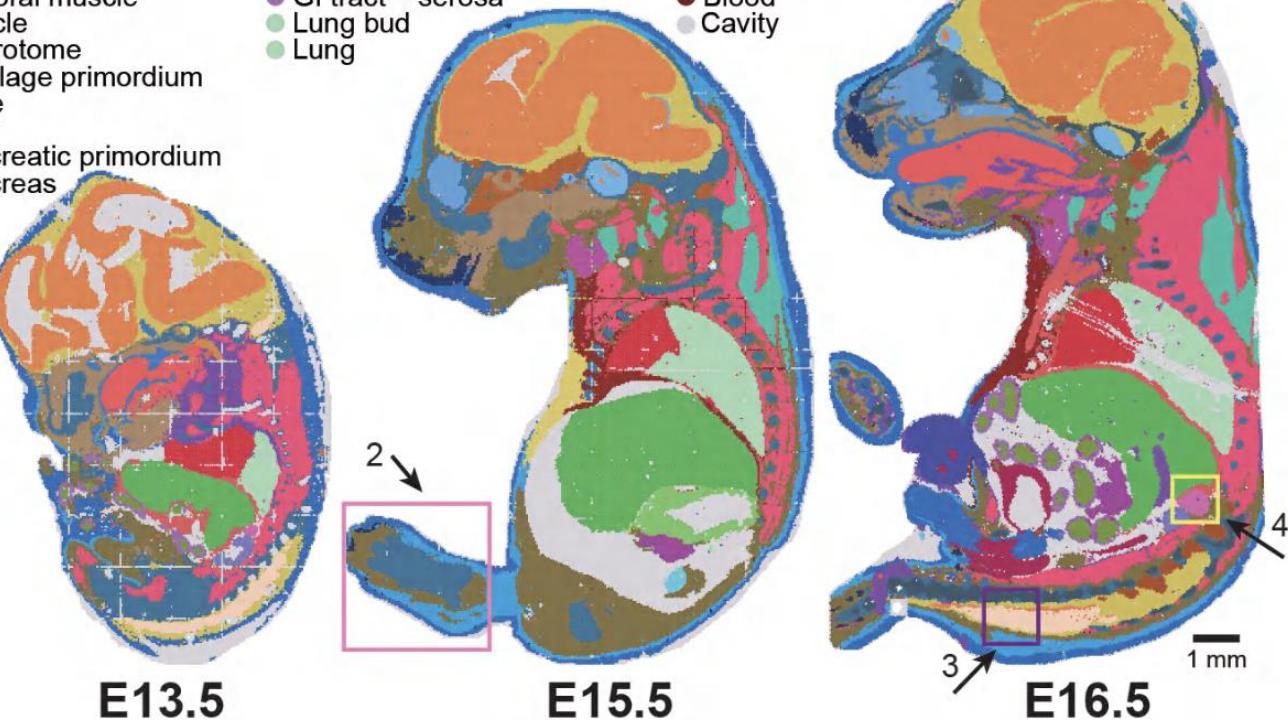


- Intermediate mesoderm
- Urogenital ridge
- Urethra
- Kidney
- Ovary
- Adrenal gland
- Salivary gland
- Abdominal wall
- Blood vessel
- Mesenchyme
- Dermis
- Connective tissue
- Adipose tissue
- Smooth muscle
- Tongue

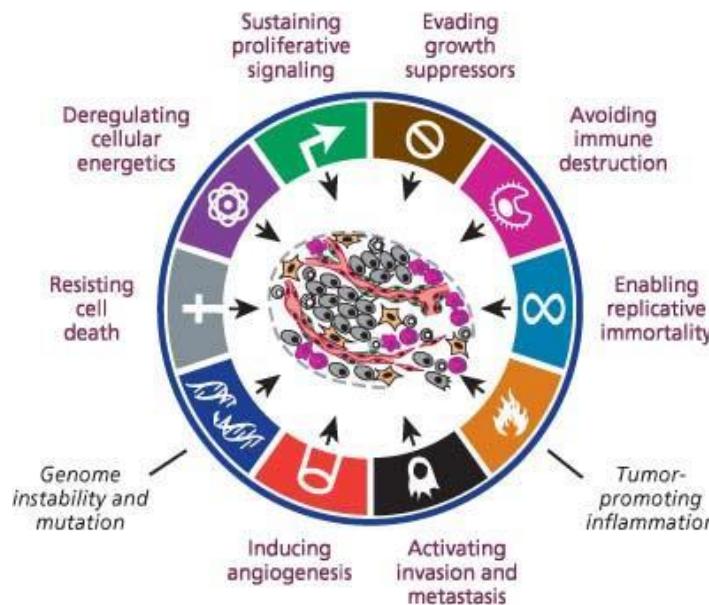
- Dermomyotome
- Diaphragm
- Pectoral muscle
- Muscle
- Sclerotome
- Cartilage primordium
- Bone
- Liver
- Pancreatic primordium
- Pancreas

- Stomach
- GI tract
- GI tract – serosa
- Lung bud
- Lung

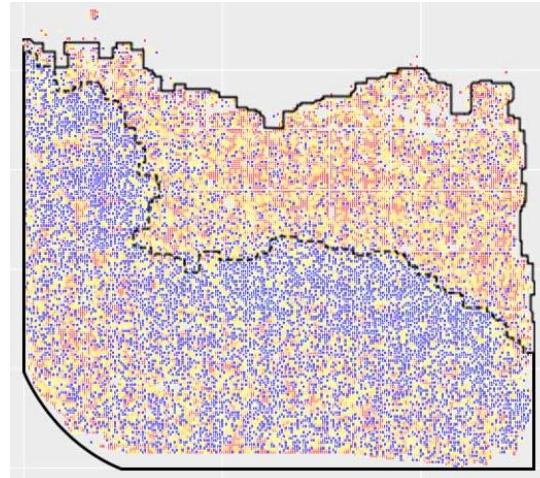
- Physiological umbilical hernia
- Umbilical cord
- Blood
- Cavity



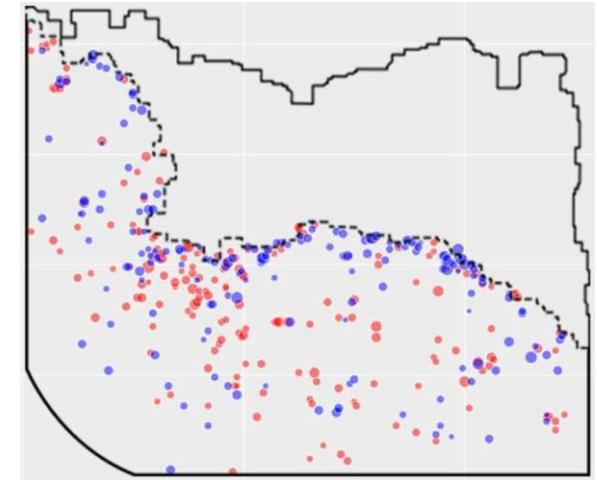
## Projection of Cancer Hallmarks to ST



细胞周期



巨噬细胞



M1

0.0

0.2

0.4

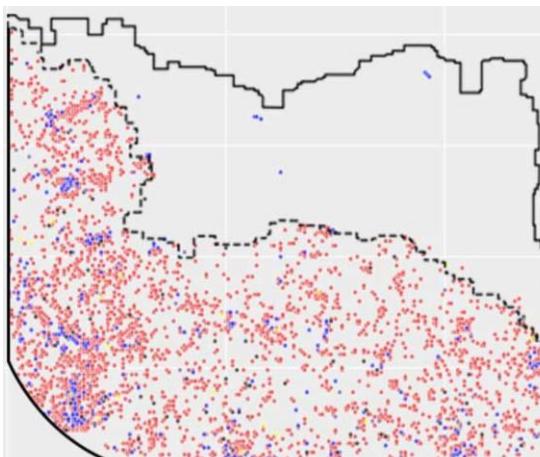
0.6

0.8

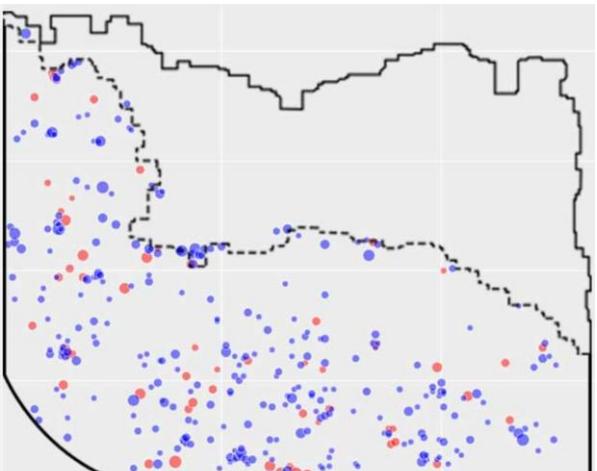
1.0

M2

成纤维细胞



T 淋巴细胞



Cytotoxicity

0.0

0.2

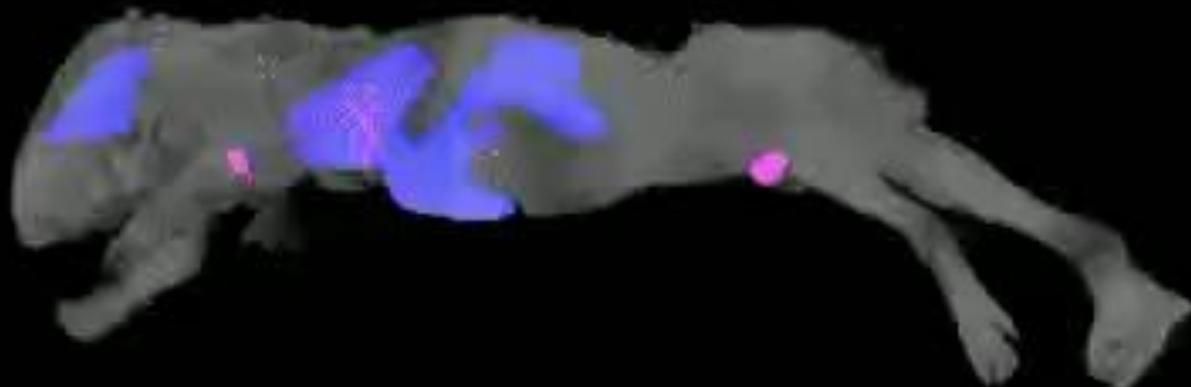
0.4

0.6

0.8

1.0

3D reconstruction of entire tumor bearing mouse by DeepMACT



5 mm



1. 基因组中为什么存在大量 “垃圾DNA” (非编码DNA) , 它们有什么作用?
2. 人类基因组中存在约40%逆转录病毒序列? 这些序列是怎么来的? 又有什么作用?
3. 决定神经细胞和心肌细胞不同的机制是什么?
4. 相同的基因组如何产生数百种不同类型的细胞, 其基本控制思想可能是什么?
4. 请设想一种能够实现 “千万级别单细胞” 通量的技术方案?
5. 假设你现在可以做一千万单细胞转录组, 你会进行哪些探索?
6. 如果通过人工手段来操控细胞的命运?
7. 请设想一种不用切片的真正3D空间转录组技术?
8. 利用时空组技术, 你想进行哪些探索, 请设计一个你感兴趣的实验?
9. 如何利用单细胞或空间组技术研究发育过程, 请设想一种实验进行探索?
10. 时空组学未来如何影响生命科学以及人类的生活?
11. 数学在生命科学的研究中起到的作用有哪些?
12. 生命科学的研究范式未来可能会发生哪些变化?

Thank you!

