

## 单细胞转录组数据获取（实验获取数据流程）

### 1. 细胞分离与制备



### 2. 单细胞捕获与裂解



### 3. 反转录与条形码标记



### 4. cDNA扩增



### 5. 文库构建与高通量测序

1.1 新鲜组织剪碎至 $1\text{ mm}^3$ ，加解离酶 $37^\circ\text{C}$ 振荡消化15 min  
1.2 终止消化，过 $40\text{ }\mu\text{m}$ 滤网，离心收集细胞  
1.3 含PI/DAPI的缓冲液染色，流式分选筛选去除死细胞，台盼蓝染色荧光计数器计数（活性 $>85\%$ ）

2.1 将细胞悬液、Gel Beads（含条形码引物）、油相注入Chromium芯片  
2.2 液滴内裂解细胞释放RNA

3.1 采用poly(T)序列的引物（结合mRNA的poly(A)尾）启动反转录，将mRNA转化为互补DNA（引物上携带有独特的分子标识符和细胞条形码）

4.1 对带有条形码和UMI的cDNA进行PCR扩增

5.1 将扩增后的、带有细胞条形码和UMI的cDNA片段化  
5.2 加上测序接头，构建成标准的二代测序文库  
5.3 高通量测序，获取组学数据

## 单细胞转录组数据获取（公共数据库资源）

### 1. 搜索单细胞数据集

- 1.1 访问GEO数据库 (<https://www.ncbi.nlm.nih.gov/geo/>) ,
- 1.2 搜索框中输入关键词 ("scRNA-seq" AND "\*\*\*")

### 2. 识别包含单细胞数据的记录

- 2.1 仔细阅读Series记录的摘要、总体设计、描述和协议部分，作者通常会明确说明这是单细胞RNA测序研究。最关键的是检查“Supplementary file”部分即下载数据的主要入口，常见格式：\*.h5、\*.h5ad、\*.loom、\*.mtx.gz、\*.csv、\*.tsv、\*.txt

### 3. 下载数据

- 3.1 下载处理过的矩阵和元数据：在“Supplementary file”部分找到需要的文件下载
- 3.2 下载原始FASTQ数据：Series 记录页，找到SRA Run Selector链接，进入选择需要的运行，点击"Accession List"下载一个包含所选运行编号的文本文件

### 4. 加载和处理数据

- 4.1 处理过的矩阵：
  - .h5: 使用Seurat::Read10X\_h5()读取；
  - .h5ad: 使用SeuratDisk::LoadH5Seurat()读取；
  - .loom: 使用Seurat::as.Seurat(loomR::connect())读取；
  - .mtx.gz+features.tsv+barcodes.tsv: 使用Seurat::Read10X读取；
- 4.2 文本矩阵：使用标准表格读取函数(read.table)；
- 4.3 元数据：使用表格读取函数加载，然后与表达矩阵关联；
- 4.4 原始FASTQ数据：使用单细胞分析流程（Cell Ranger）进行比对、定量，生成计数矩阵后才能进行下游分析

# 常见问题解析

## 1.单细胞样本质量低有哪些原因？

①样本来源不佳，样本来自手术切除的陈旧组织、尸检样本、冷冻保存不当的组织，细胞天然活性差。②酶解时间过长、温度过高、酶浓度过高，物理研磨过度，导致细胞膜破裂、RNA降解，未完全解离成单细胞悬液，残留细胞团块堵塞仪器或导致多细胞捕获；③冻存过程或复苏过程不当，导致大量细胞死亡；④细胞浓度过高易导致双胞胎/多胞体，浓度过低则捕获效率低下。

## 参考文献

- [1]Wang X, He Y, Zhang Q, et al. Direct Comparative Analyses of 10X Genomics Chromium and Smart-seq2 [J]. Genomics Proteomics Bioinformatics, 2021, 19(2): 253-266.
- [2]Song L, Li K, Hong X, et al. Transcriptomic evidence of lung repair in paediatric ARDS survival [J]. Clin Transl Med, 2023, 13(8): e1366.