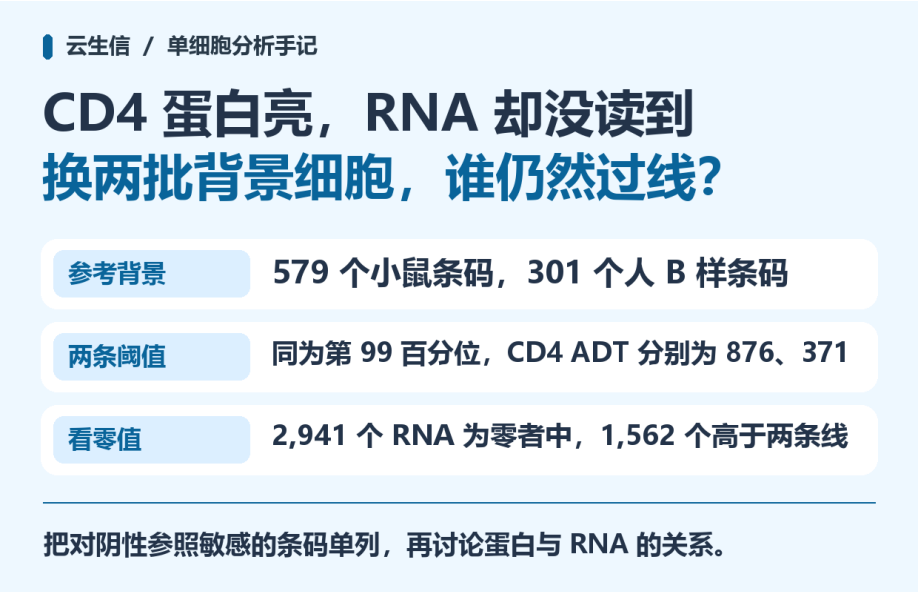
CITE-seq 里 CD4 蛋白阳性、RNA 却是零，怎样判断这个细胞?

在同一颗公开 CITE-seq 细胞上核对 CD4 RNA 和抗体原始计数，分别用小鼠掺入细胞和人 B 样细胞估计背景，区分两种背景都支持与只在一种门槛下阳性的条码。



运行教学示例

Python 3.10 或以上及 requirements.txt 可离线复核。原矩阵的流式提取脚本和校验值随包。

```
python run.py --input example --out results
```

在解压后的资料包根目录运行。example 是教学输入，results 保存本次输出，expected 保存可供核对的参考结果。

输入字段

paired_counts.tsv、source_hashes.json 真实的 RNA、ADT、物种分子数与条码，全部 8,617 行。

rebuild_cite.py 从发布的 RNA 和 ADT 文件重新提取同条码数据。

核对输出，再换自己的数据

background_cuts.tsv、rna_background_cross.tsv 两个独立定义的参照群、阈值和 RNA 检出交叉表。

barcode_evidence.tsv、example_barcodes.tsv、background_and_joint.png 完整条码判定、可回查实例、背景分布及 RNA/ADT 散点。

用示例验算

1 对应 579 个小鼠参照与 301 个人 B 样参照，第 99 百分位分别为 876 与 371。

2 3,395 个 T 候选中 2,941 个未检出 CD4 RNA，1,562 个超过双线、956 个依赖参照、423 个两线均未超过。

3 454 个检出 RNA 的条码里仍有 14 个未高于两条线。

换用真实数据

按样本选择有实验依据的负参照，保留其原始 ADT 分布。把 RNA 与蛋白按同一条码对齐，量化对背景选择的敏感度。

小鼠和人 B 样条码都有局限，两者都不能替代匹配的技术背景、等型对照或独立流式测量；CD4 抗体与 RNA 的零值不构成谱系判决。

方法资料

CITE-seq 原始论文

作者数据的 Seurat 示例