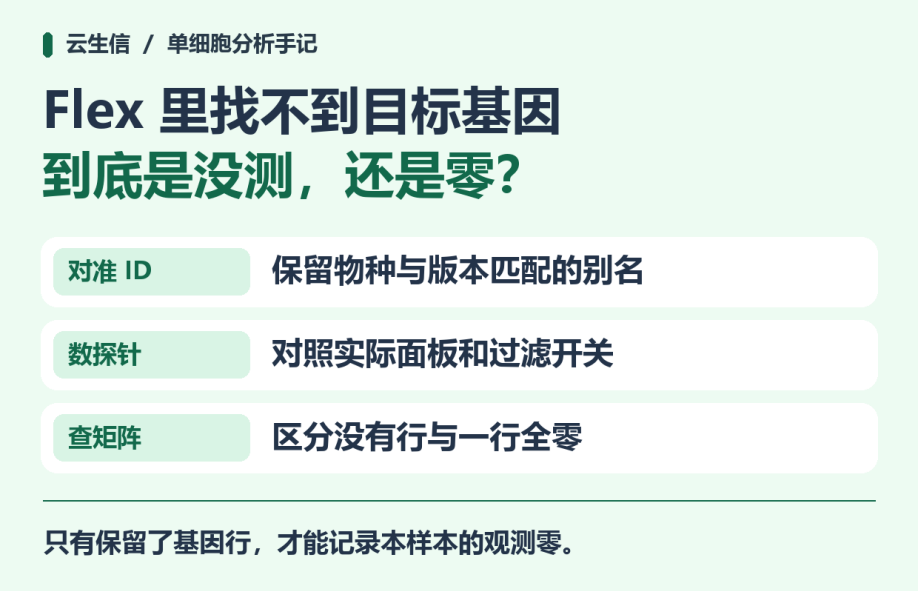
10x Flex 矩阵里完全找不到目标基因，该从哪里查起?

把一个查不到的目标基因沿着物种注释、实际 Flex 探针集、过滤参数和样本特征矩阵查到底，给每个目标留下可复核的原因和原始计数。



运行教学示例

Python 3.10 或以上，需 numpy 和 scipy，可通过 requirements.txt 安装。

```
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

probe_set.csv、run.json 原始面板 CSV 和实际运行的样本、物种、面板、软件及 filter-probes 设置。

targets.txt、aliases.tsv、filtered/ 目标清单、同版本名称与 ID 对照表、该样本未经二次筛选的过滤矩阵。

核对输出，再换自己的数据

targets-audit.tsv 逐基因列出 ID、探针数、included 数、矩阵有无该行、观测总量和判定。

sample-run.json 保留用于解读这张表的参数以及 RNA 特征数与样本条码数。

用示例验算

1 ALPHA 有两条纳入探针，样本里共测到 6 个 UMI。

2 BETA 存在于矩阵中，两个条码合计为零。GAMMA 的两条探针只有一条标为 TRUE，按默认探针过滤从面板剔除。

3 EPS 有合格探针但所给矩阵缺行，列为待查。OLDX 对应两个 ID，暂缓归属。

换用真实数据

核实实验物种及试剂版本，将 Cell Ranger 输出根目录复制的 probe_set.csv 和对应样本 sample_filtered_feature_bc_matrix 放入练习目录。过滤开关取自本次 multi 配置，不要用官网其他版本的面板替换实际使用文件。

未覆盖、被过滤和零计数属于不同状态。脚本只校验提供的面板、别名表和样本矩阵，并不会依据名称猜测基因，也不能从一个样本的零值断定组织中不存在该转录本。

方法资料

10x Flex 探针集字段和过滤规则

Cell Ranger Flex 运行输出