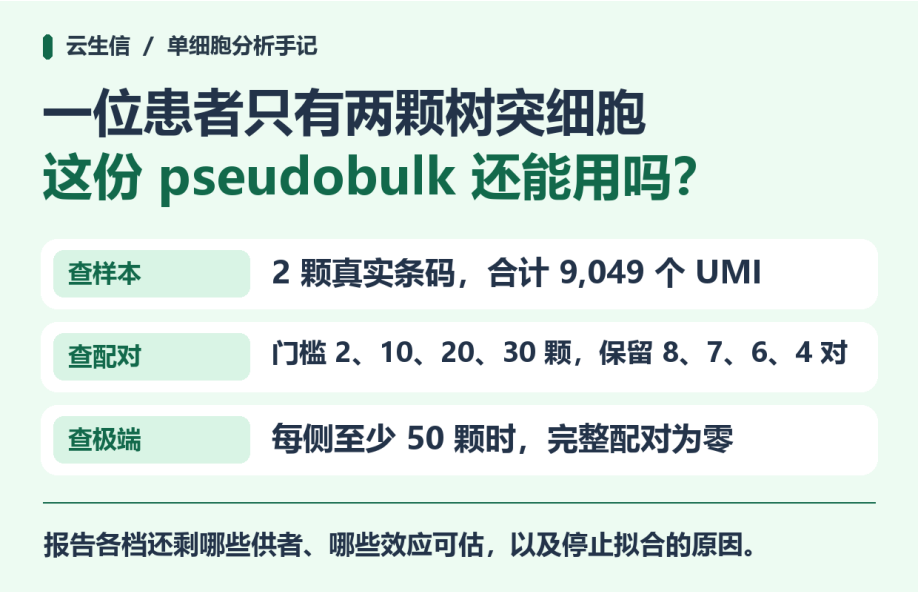
某位患者只有 2 个树突状细胞，这个 pseudobulk 样本怎么处理?

用真实树突状细胞矩阵，把一位供者的对照样本从十颗可追踪地缩减到两颗，比较六档门槛下的样本数、完整配对数与 edgeR 效应。



运行教学示例

Python 3.10 或以上与 requirements.txt。完整重新拟合需要 R、edgeR、Matrix、statmod；四档通过覆盖要求的模型已实际运行。

```
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

counts.mtx.gz、genes.txt、cells.tsv 529 颗带供者和条件的真实树突状细胞，一万五千七百零六个基因。

fit_gates.R、example/fit/ 受控缩减条码、六档覆盖门槛、完整配对、共同基因集和已拟合结果。

核对输出，再换自己的数据

coverage.tsv、included_depth.tsv 对照侧、刺激侧与两侧齐全的供者数，每份样本纳入的细胞及计数。

focus_genes.tsv、coverage_and_effect.png 四档可比较模型的 ISG15、IFIT1 效应和 FDR，以及随门槛变化的图。

用示例验算

1 受控缩减后目标侧 2 颗计 9,049 UMI，配对刺激侧 18 颗、55,754 UMI。

2 六档保留 8、7、6、4、1、0 个完整供者对。

3 前四档使用同一批 1,382 个基因，ISG15 的 log2FC 在 6.58 至 6.76 之间。

换用真实数据

保留原始样本及细胞类型清单，先写出纳入和配对规则；画细胞数及总计数，对预设的数档覆盖门槛汇报剩余供者、设计和模型效应。

两颗是教学干预构造出的低覆盖样本；本次的门槛和拟合条件用于展示判断过程，不足以界定其他研究的最佳门槛。

方法资料

Kang 等的刺激研究与供者设计

muscat 按细胞类型和样本检查覆盖