稀有细胞的开放峰在全样本里看不见，按细胞类型找峰能找回来吗?

用同一份带条码的模拟 ATAC 片段实际运行 MACS3，比较全样本、常见群和稀有群的峰，再按供者回查每条曲线的细胞数。



运行教学示例

Python 3.10 或以上，安装 requirements.txt。重跑找峰另安装 MACS3，参考版本 3.1.0。

```
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

fragments.tsv.gz、cells.tsv、loci.tsv 三名供者的模拟片段、条码标签及三个测试区间。

call_peaks.py、fit/*.narrowPeak 可改组别的找峰脚本与 MACS3 真正产生的候选。

核对输出，再换自己的数据

peak_calls.tsv、donor_support.tsv 逐区间的找峰结果、每位供者的片段与条码支持。

coverage.tsv、donor_coverage.png 四十碱基分箱的原始和标准化覆盖。

用示例验算

1 中间候选在稀有群有一个峰，合并与常见群均为零。

2 稀有候选在三位供者里各有六个条码贡献片段。

3 另一个候选的稀有群片段为 61、0、0。

换用真实数据

将同组细胞按生物学重复留档，从同一版 fragments、同一基因组和同一套调用参数开始。固定候选边界后逐供者重数。

模拟结果验证峰调用和支持表的关系，不能直接说明真实人类基因组某个位点的功能。

方法资料

Signac 按群找峰

MACS3 成对片段格式和参数