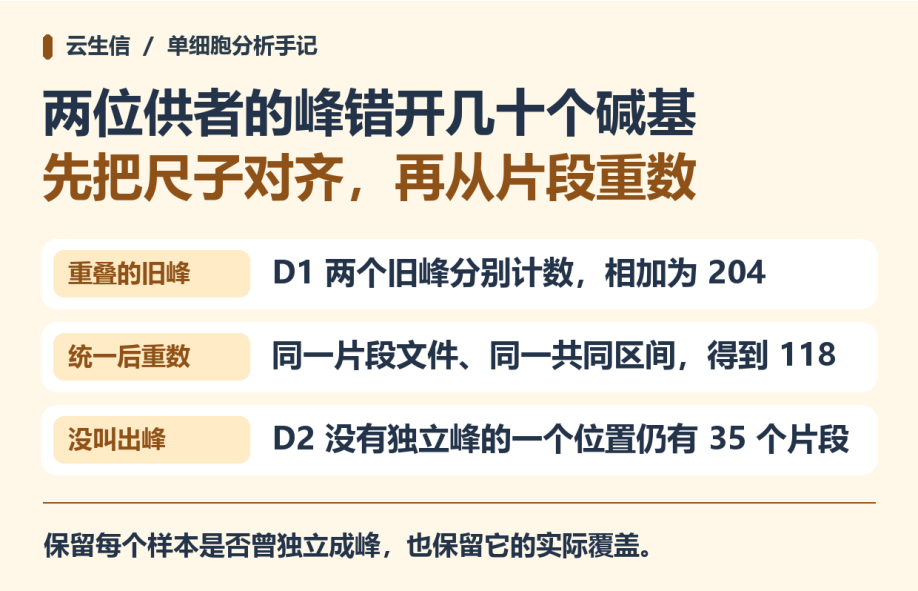
两个供者的 ATAC 峰边界差了几十个碱基，还能比较开放程度吗?

在两份带独立峰边界的片段文件上，直接比较旧计数相加和共同区间重计数，保留跨边界的反例。



运行教学示例

Python 3.10 或以上，先安装 requirements.txt。

```
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

D1.bed、D2.bed、manifest.tsv 两套原始峰、参考组装与坐标声明。

两份 fragments.tsv.gz、make_example.py 按供者保存的完整去重片段，以及可编辑的生成器。

核对输出，再换自己的数据

common_peaks.bed、common_counts.tsv 共同区间、分供者的记录数及条码数和深度分母。

before_after.tsv、same_interval_counts.png 与各样本原峰之和对照，标记原样本未被调用的区间。

用示例验算

1 九个旧峰合成六个共同区间；刚好相接的两段仍分别保留。

2 第一处 D1 旧计数和为 204，重数为 118；D2 则由 108 变为 115。

3 D2 在未单独成峰的 [4000,4100) 区间确有 35 个重叠片段。

换用真实数据

先统一参考组装和 BED 坐标，再根据项目选择合并或固定宽度。保留原始片段并按同一组共同区间重新计数。

同供者的多个细胞不能代替生物学重复；共同区间较宽时可能糅合相邻的独立调控区。

方法资料

Signac 分组找峰

BEDTools 区间相交定义