一个开放峰就在目标基因旁边，凭什么把它列为候选增强子?

从同滴配对的 1,050 个外周血细胞出发，逐个比较 MS4A1 附近十五个峰，查询可比背景与真实片段切口。



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

counts.npz、peaks.tsv、cells.tsv 实际峰计数、坐标、MS4A1 与独立谱系标记、两种全深度。

locus_fragments.tsv.gz 公共原始片段中与 30 万碱基窗口重叠且属于所选条码的记录。

核对输出，再换自己的数据

associations.tsv、matched_background.tsv 两套人群中的相关和每个峰的五十个匹配背景。

fragment_cut_tracks.tsv、peak_coverage.tsv、两张 PNG 真实片段切口轨迹、逐峰强度和分层相关。

用示例验算

1 第一近端示例在混合组相关约 0.5177，B 样群内约 0.1354。

2 共 15 个局部峰与 4,000 个候选背景，每个局部峰各匹配 50 个。

3 三组真实片段在 250 碱基窗口上计数，各有 1,200 箱。

换用真实数据

先换同一版本的基因坐标、独立标记和条码，检查两组真实深度。另加入 GC 匹配及独立供者验证后，再选择需要扰动的区间。

本例透明匹配检出率、平均计数和长度，没有加入 GC 或验证调控方向；经验尾部是候选排序的辅助信息。

方法资料

Signac 的 LinkPeaks 背景模型

10x 公开真实配对数据