



基于流式细胞荧光分选的 孔板scATAC-seq 测序方法的 标准化分析流程开发

完成人： 卢绍荣

指导老师： 徐彦辉

1

课题背景

- Plate-based scATAC
- 测序数据分析

2

开发框架

- WDL
- 依赖软件
- 运行框架

3

开发结果

- 分析结果
- 分析报告
- 项目获取

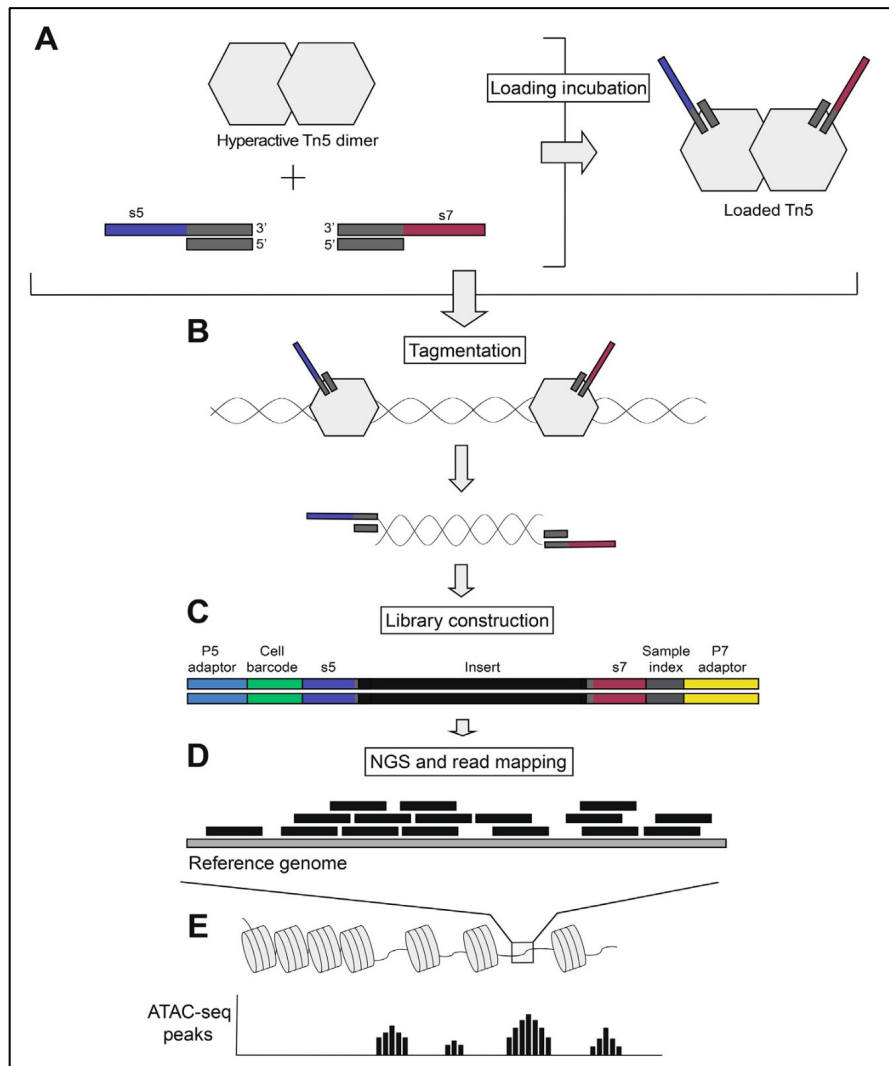


课题背景



孔板scATAC-Seq技术

经典 scATAC 流程



孔板scATAC流程

Tn5 改造

Tn5
附着切割

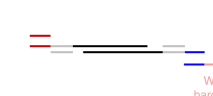
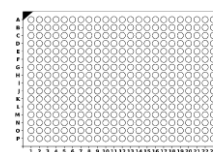
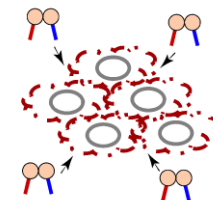
文库构建

下游分析

附着

流式
分选

切割



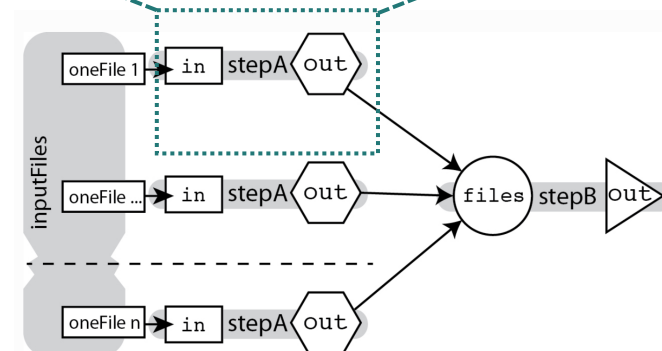
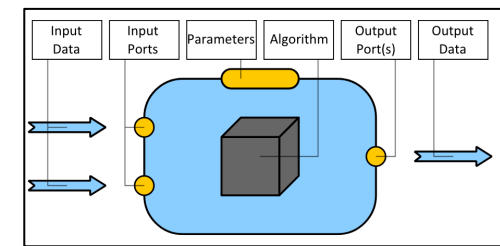
scATAC-Seq数据分析

非科学问题依赖

- 修整和剪切
- 序列联配
- 去PCR重复
- Call peaks
- Black list 去除
- 质量控制
- 生成特征矩阵
-

可并行多线程

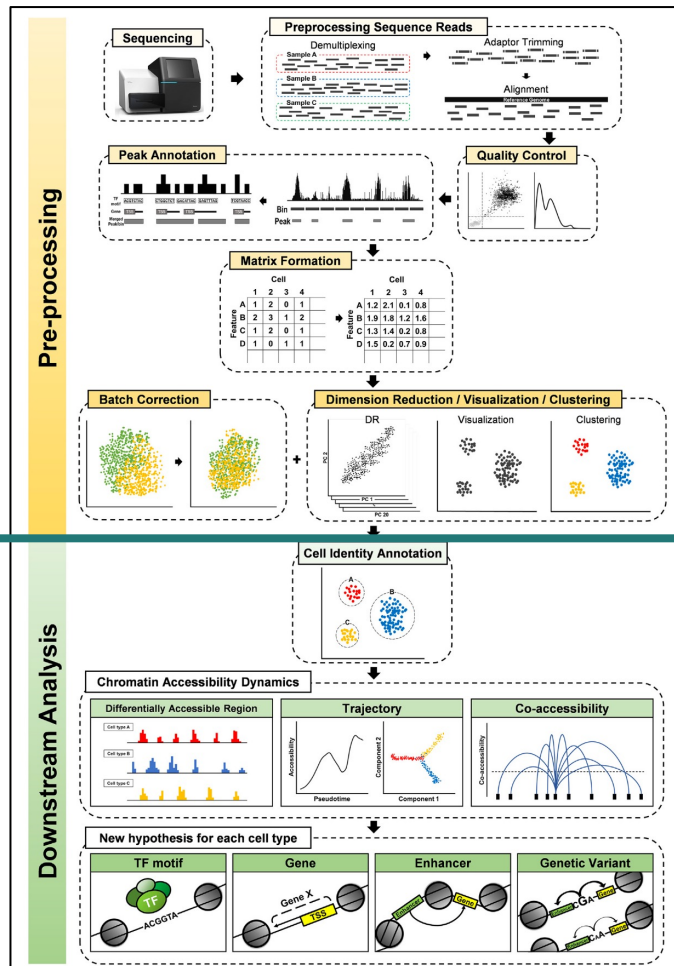
组织



科学工作流管理系统

- 并行处理
- 多任务分配
- 资源最大化
- 环境隔离

移植优化、性能优化



scATAC 分析流程

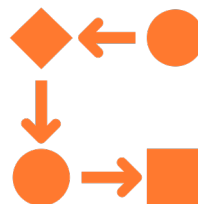


开发框架



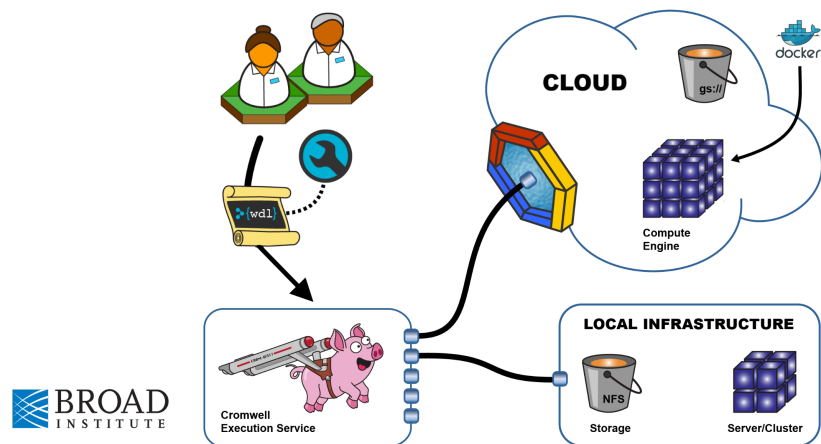
搭建体系

snATAC-wdl



框架

Meet WDL + Cromwell



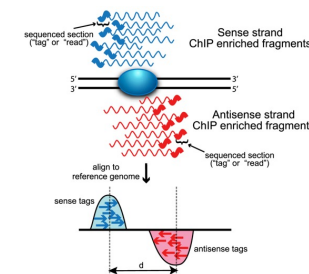
BROAD INSTITUTE

Working Description Language (WDL)



构成

- Cutadapt
- Fastp
- Bowtie 2
- Picard
- Samtools
- Bedtools
- MACS2
-



Bowtie 2
Fast and sensitive read alignment

管线框架

• 先分

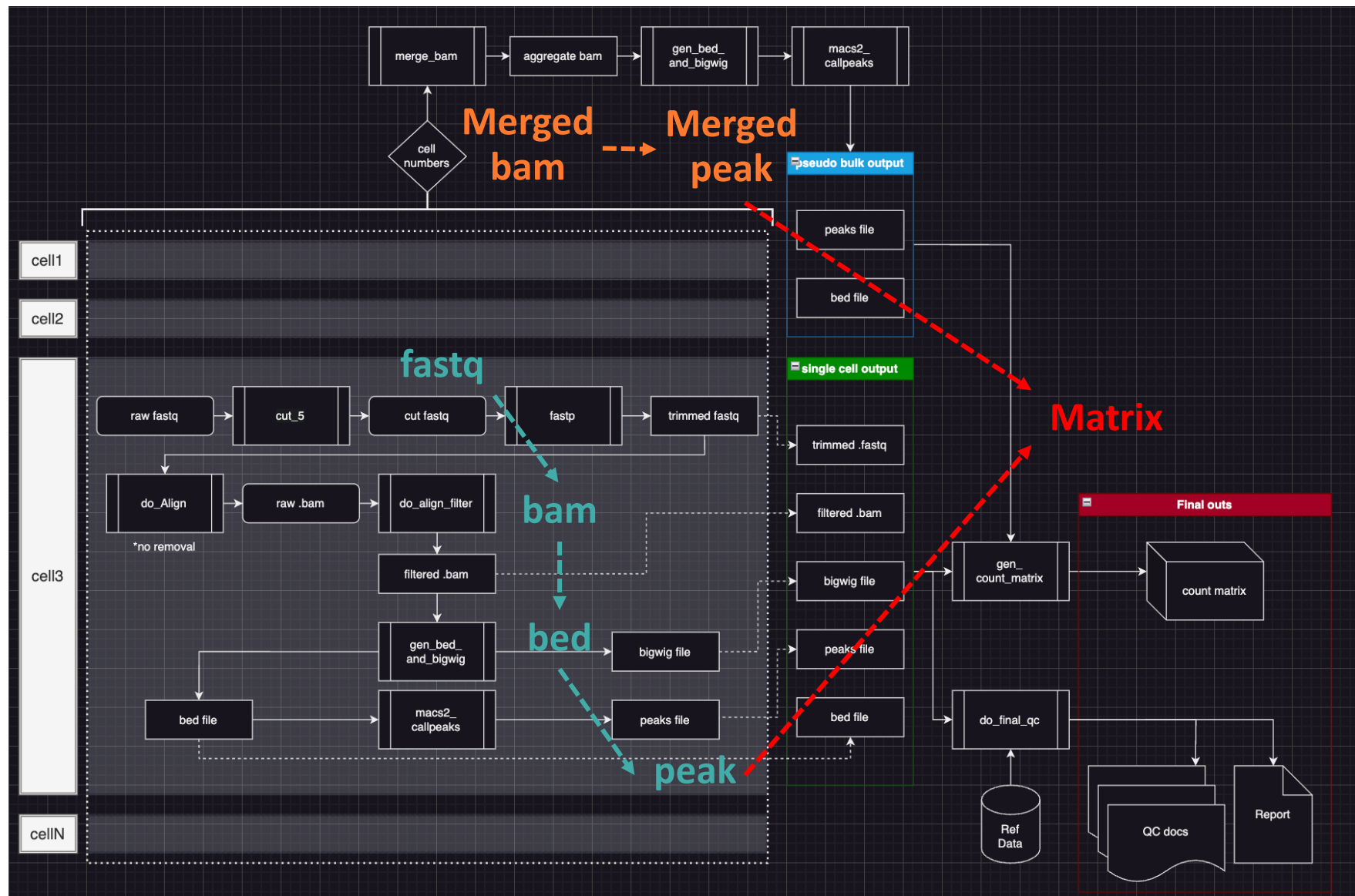
• 单细胞

- ✓ 切除
- ✓ 过滤
- ✓ 比对
- ✓ 去重
- ✓ 压缩
- ✓ 转换
- ✓ Callpeaks

• 后总

• Pseudo bulk

- ✓ 总和
- ✓ Callpeaks
- ✓ 计算
- ✓ QC





开发结果

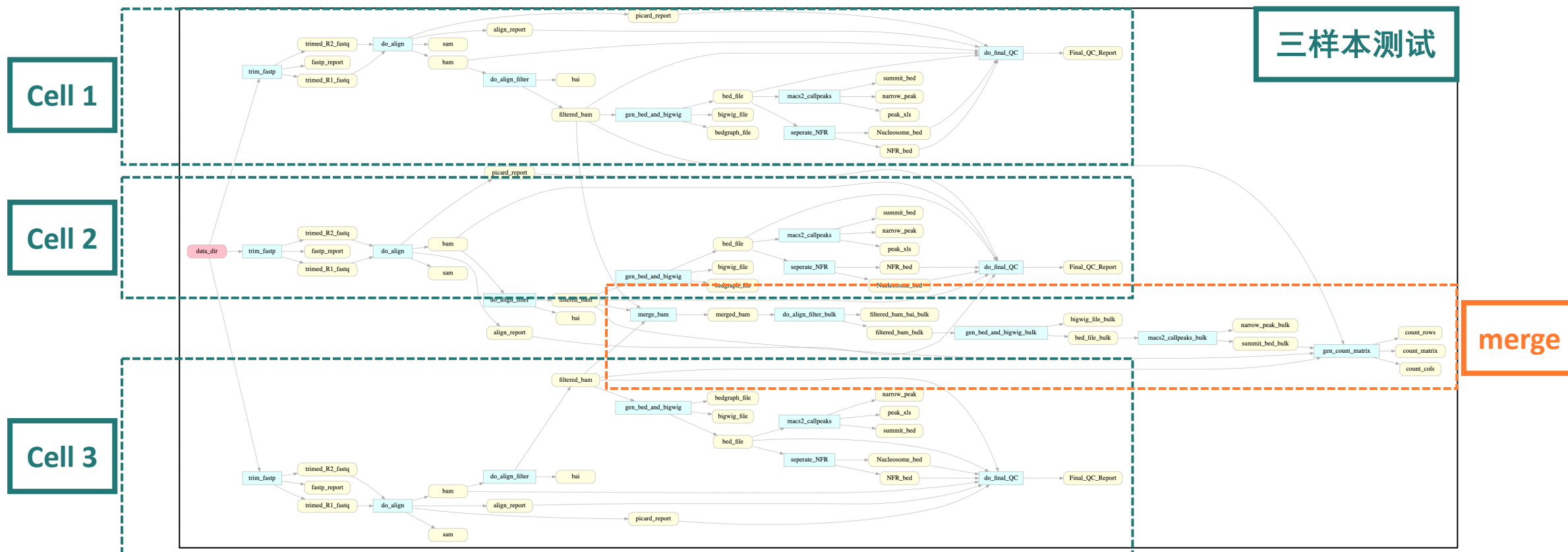


运行结果

```
(work) [lushaorong gpu19 ~/work/snATAC-release/hskin_test/out_sub/croo] WORK 0 >> tree -L 3
.
├── croo.filetable.f5fa87dc-d7cc-447a-9ae8-c107dc9fcf29.tsv
├── croo.report.f5fa87dc-d7cc-447a-9ae8-c107dc9fcf29.html
├── Files
│   ├── CountMatrix
│   │   ├── hs_skin_count.cols -> /work/xulab/lushaorong_work/snATAC-release/hskin_test/out_sub/out/snATAC/f5fa87dc-d7cc-447a-9ae8-c107dc9fcf29/call-gen_count_matr
│   │   └── ix/execution/glob-4aedce1bb784667581c296c22b32cab1/hs_skin_count.cols
│   │       ├── hs_skin_count.mtx -> /work/xulab/lushaorong_work/snATAC-release/hskin_test/out_sub/out/snATAC/f5fa87dc-d7cc-447a-9ae8-c107dc9fcf29/call-gen_count_matri
│   │       └── ix/execution/glob-f988005ddae426e510b256b8bdcd44d2/hs_skin_count.mtx
│   │           ├── hs_skin_count.rows -> /work/xulab/lushaorong_work/snATAC-release/hskin_test/out_sub/out/snATAC/f5fa87dc-d7cc-447a-9ae8-c107dc9fcf29/call-gen_count_matri
│   │           └── ix/execution/glob-b428827dbc53a3ae687e272a72217156/hs_skin_count.rows
│   ├── pseudo_bulk
│   │   ├── Align
│   │   ├── Bed
│   │   └── Peak
│   └── sn
│       ├── Align
│       ├── Bed
│       ├── Fastq
│       ├── Peak
│       └── Reports
```

运行输出

- ✓ 单细胞
- ✓ 多细胞
- ✓ 重要中间文件
- ✓ QC报告
- ✓ 特征矩阵

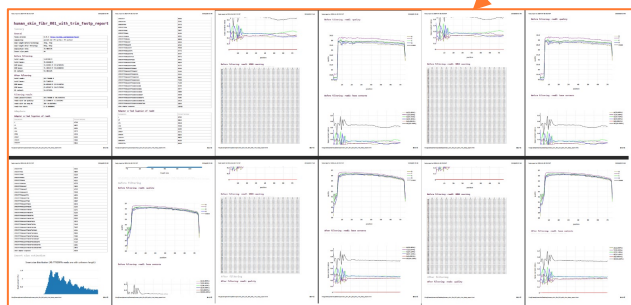




质量报告

Part1:质量报告

- ✓ Adapter情况
- ✓ 测序质量
- ✓ GC分布
- ✓ 预测片段长度
- ✓ 附完整报告

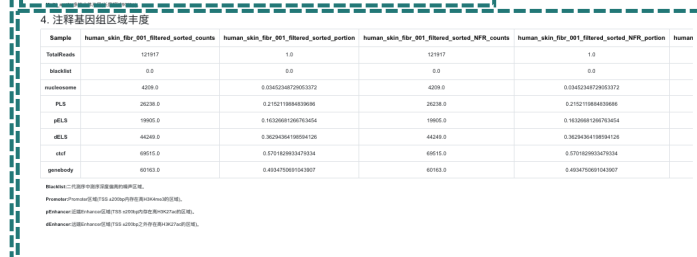
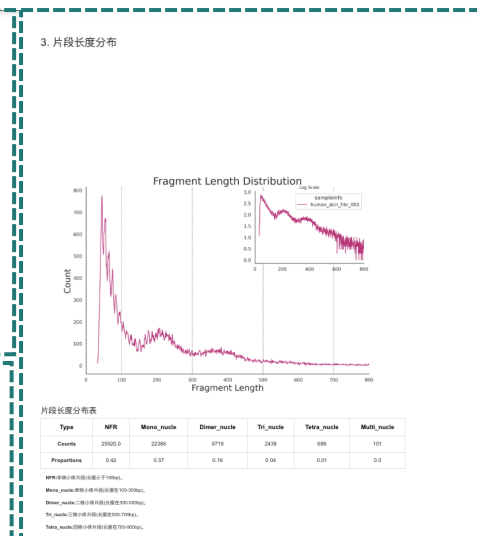


Part2:比对报告

- ✓ 总reads数
- ✓ 比对率
- ✓ 正确比对率
- ✓ chrM比率
- ✓ 重复率
- ✓ 测序饱和度曲线

Part3:片段长度

- ✓ 片段长度分布图
- ✓ 片段长度比率



Part4:片段注释

- ✓ 已注释区域富集情况
- ✓ Blacklist、nucleosome、ctcf、genebody.....

项目获取



<https://github.com/Phoenix-Lu/snATAC-wdl>

snATAC-wdl Public

main 2 Branches 0 Tags

Go to file Add file Code

Phoenix-Lu Add files via upload 8e843e6 · 2 weeks ago 14 Commits

File/Folder	Action	Time
example_QC_reports	Add files via upload	2 weeks ago
example_croo_out	Add files via upload	2 weeks ago
src	Add files via upload	2 weeks ago
wdl_profile	Add files via upload	2 weeks ago
ATAC.yaml	Add files via upload	2 weeks ago
README.md	Add files via upload	2 weeks ago

README

snATAC-wdl

Step1: get dir

```
cp /work/xulab/xulab-seq/lxr/PhoenixPipeline/snATAC-release YOURPATH
or
clone this repository it from github
```

运行示例

所有源码

环境配置

ReadMe文档

- ✓ 能使用的管线
- ✓ 完整的质控
- ✓ 详细功能模块



欢迎老师批评指正

基于流式细胞荧光分选的孔板scATAC-seq 测序方法的
标准化分析流程开发