WP3

Approach

- Obtain ATAC data per cluster

  → .bam file with all cells + cell assignment table from WP 1+2

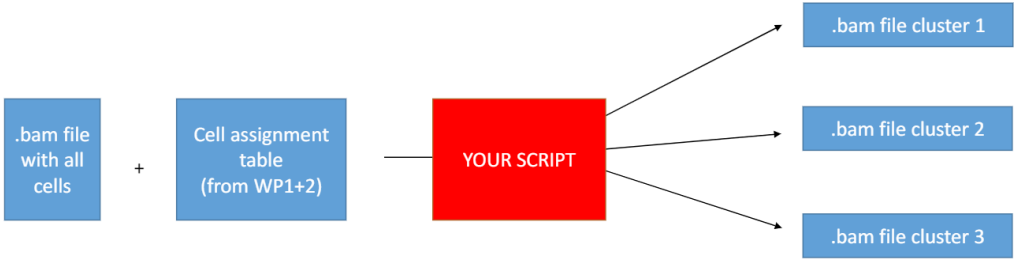
  → bam files for each cluster
- Run TOBIAS footprinting for each cluster
- Comparison between clusters + normalization

Splitting the .bam file into clusters

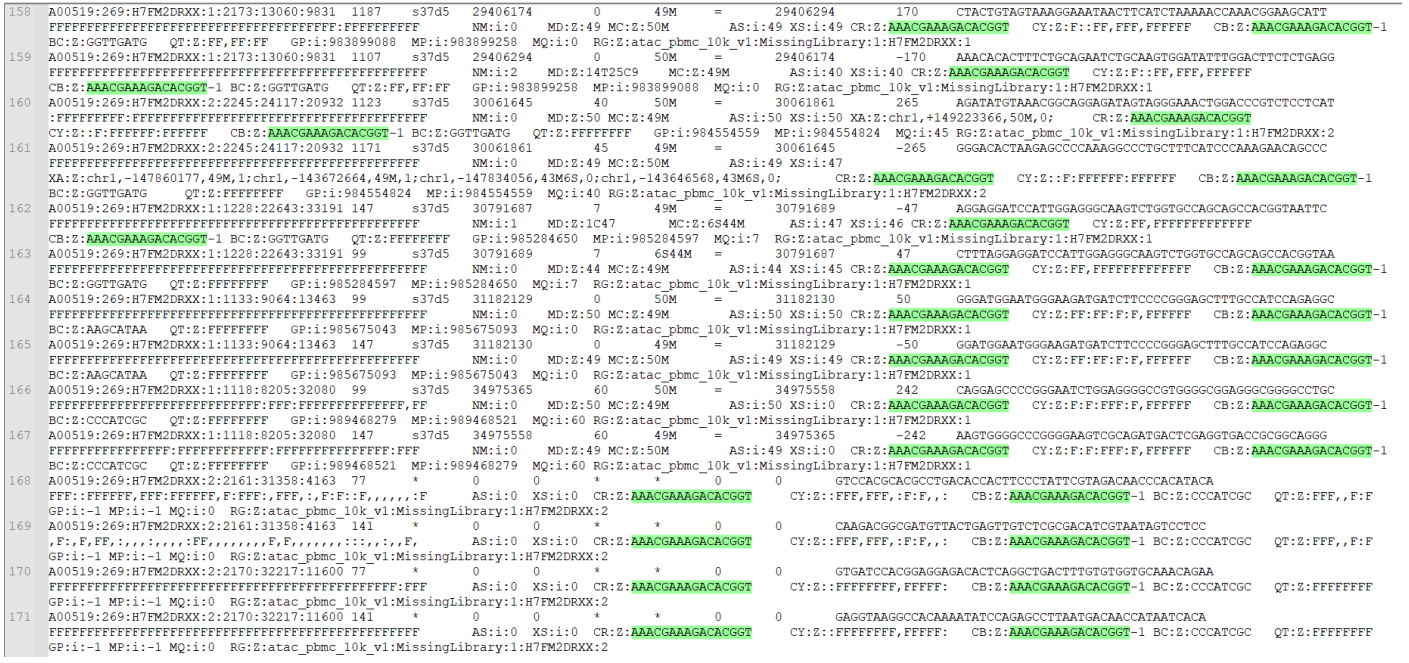
- Reads have barcodes
- Script to divide the file based on these bar codes and the assignment table
- Input:

  - ATAC of all cells (.bam)

  - barcodes and clusters (.tsv)
- Output: one .bam file per cluster



Test data



```
Barcode Cluster
AAACGAAAGACACGGT-1      2
AAACGAAAGAGGTGGG-1      10
AAACGAAAGCACGTAG-1      14
AAACGAAAGCGCCTAC-1      3
AAACGAAAGCTTTCCC-1      2
AAACGAAAGGCGTCCT-1      1
AAACGAAAGGCTTTAC-1      4
AAACGAAAGTGATATG-1      14
AAACGAAAGTTTGGAA-1      2
AAACGAACAAACGACG-1      2
AAACGAACAATTGCCA-1      4
AAACGAACACCTCGTT-1      12
AAACGAACACGGTGTC-1      14
AAACGAACATCAGAAA-1      12
AAACGAAGTCAACTGT-1      6
AAACGAAGTGATTCCA-1      10
AAACGAAGTTTGACCA-1      11
AAACGAATCCAACGCG-1      2
AAACGAATCCATCTAT-1      2
AAACGAATCCCTAAAG-1      3
AAACGAATCCTCAAGA-1      11
AAACGAATCGCATAAC-1      11
AAACGAATCGGGACAA-1      9
AAACGAATCTGAACGT-1      2
AAACGAATCTGTCGTC-1      7
AAACTCGAGAAAGCAG-1      2
AAACTCGAGAGCTCCC-1      13
AAACTCGAGAGTGGTC-1      7
AAACTCGAGCGAATAT-1      4
AAACTCGAGCGTGTTT-1      9
AAACTCGAGGTACATA-1      2
AAACTCGAGTCGCGTC-1      4
AAACTCGAGTCTTAGC-1      4
AAACTCGAGTTCAAGA-1      12
```

Running TOBIAS

- Requirements
  - .bam files of each cluster
  - Depending on tool:
    - reference genome (.fasta)
    - peaks (.bed)
    - motifs (.pfm / .jaspar / .meme)
    - output of previous tools in the pipeline
- Main output: footprint score → heatmap to compare clusters