Swiss Institute of Bioinformatics

INTRODUCTION TO NGS ATAC-SEQ DATA ANALYSIS

ATAC-seq Introduction

Dr. Joana Carlevaro Fita

October 15-16, 2025

elixir SWITZERLAND

Learning objectives

1. Introduction to the study of open chromatin accessibility from a biological point of view

2. Introduction to ATAC-seq technology

    - What type of analysis allows?

    - What requirements / experimental design is needed?

3. ATAC-seq analysis workflow overview

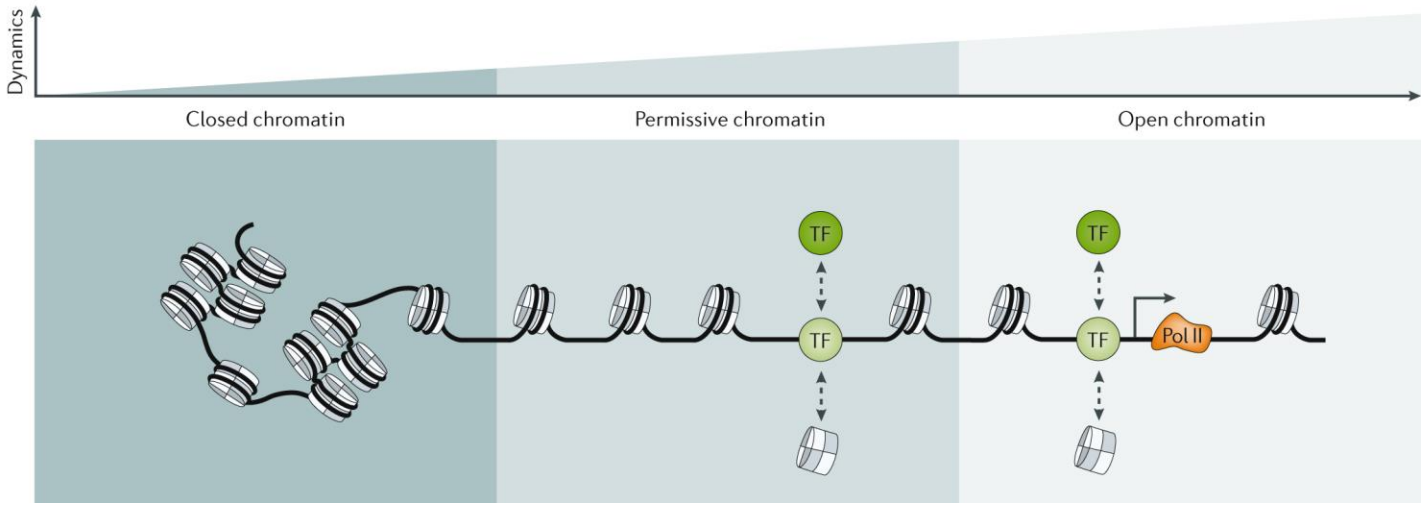
ATACseq

WHAT DOES ATAC-SEQ MEASURE?

ATAC-seq: Assay for Transposase-Accessible Chromatin using sequencing

ATAC-seq is a method used for profiling accessible chromatin across the genome

Transcription: Inactive state Active state



The nucleosome — a core structural element of chromatin — consists of an octamer of histone proteins encircled by ~147 bp of DNA

Adapted from: Klemm SL, Shipony Z and Greenleaf WJ, Nature reviews genetics, 2019

ATAC-seq

Topological organization of nucleosomes is non-uniformly distributed

Depleted at regulatory loci like:

- Enchancers and promoters
- Insulatores
- Transcribed gene bodies

→ Higher chromatin accessibility

ATACseq is used to profile chromatin accessibility and study regions under regulation vs regions being silent

Tn5 transposase

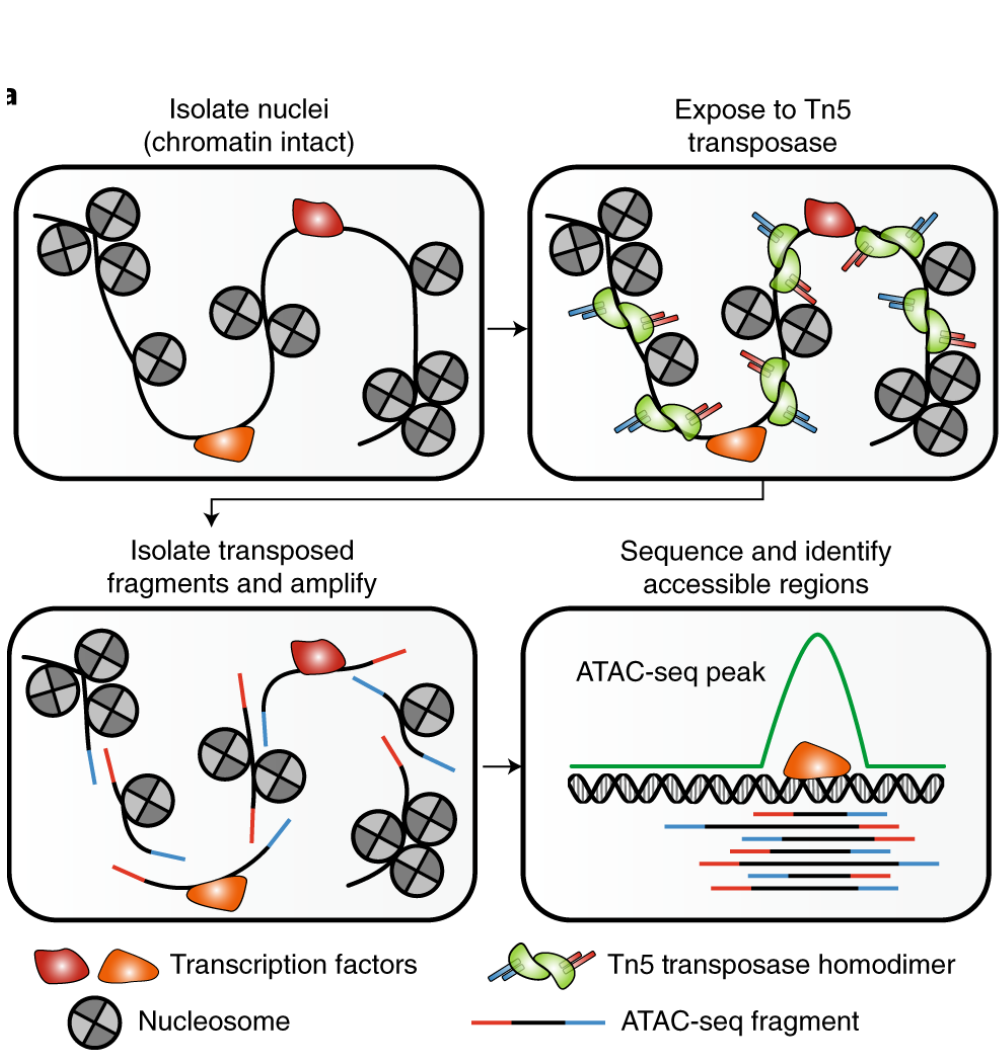
Tn5 is a hyperactive transposase, a DNA-modifying enzyme discovered in the early 1970, while investigating transposons in bacteria.

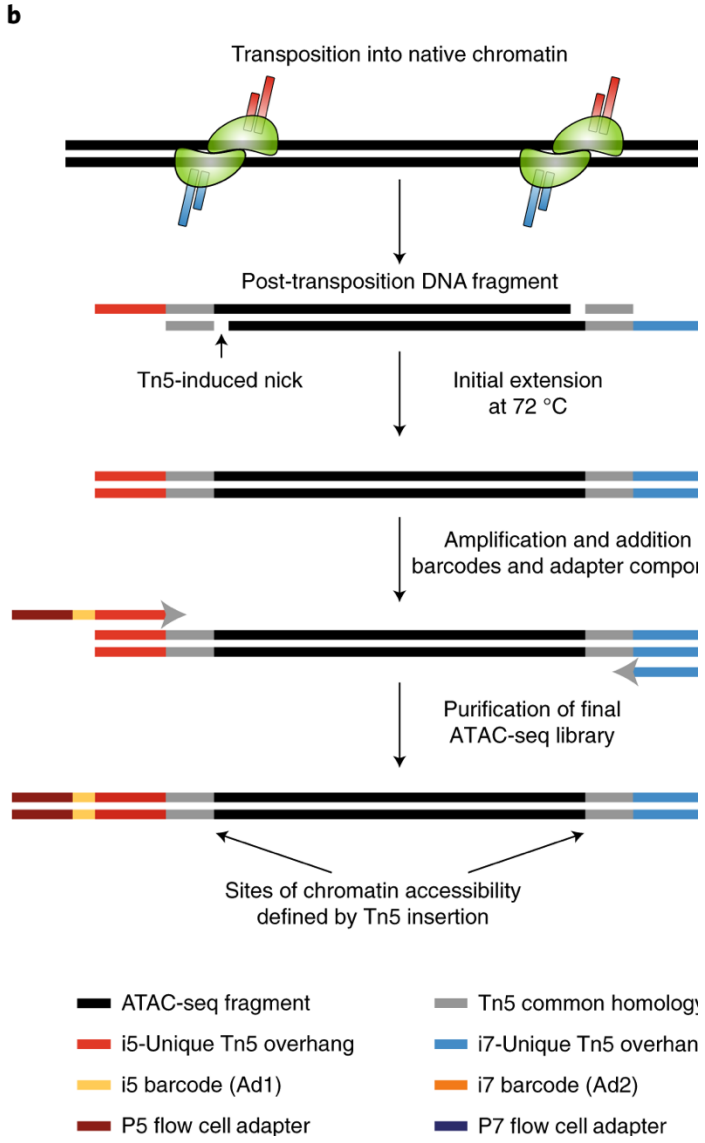
Tn5 transposase was discevered in E.Coli.

Transposon cabable of moving and inserting itself into the genome of other bacteria.

Tn5 can be loaded in vitro with DNA sequencing adapters, bread DNA (fragmentation) and ligate a sequence.

ATAC-seq technology overview



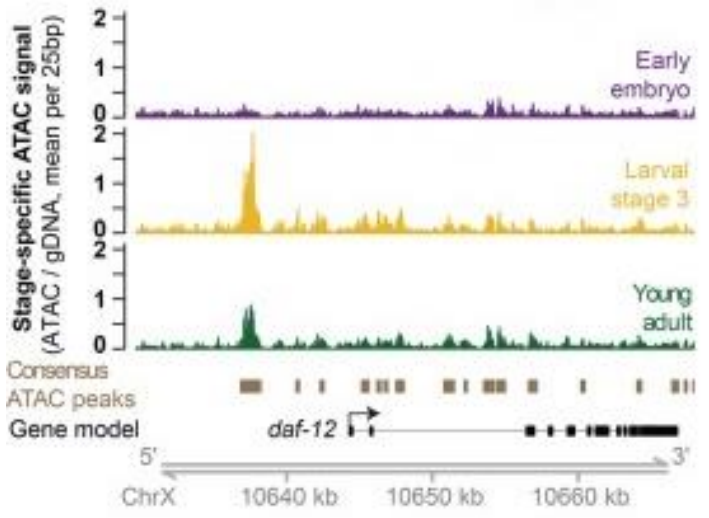


Adapted from: Grandi F. et al, Nature Protocols; 2022

ATAC-seq applications

Main applications

- Identify open chromatin regions
- Compare chromatin accessibility between conditions
  - ATAC-seq signal at given genomic regions
  - Differential Accessibility (DA) regions
- Predict TF binding
  - Motif centric: motifs trained by Chip-seq or DNAse-seq
  - De novo discovery: known or new motifs
- Map nucleosome positions



ATACseq is used to profile chromatin accessibility and compare epigenetic landscapes between cell lines

Daugherty AC et al, Europe PMC; 2017

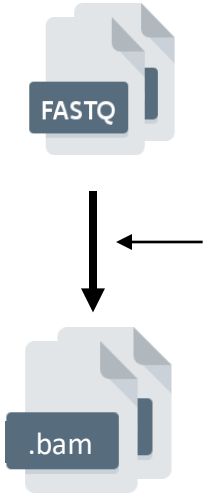
ATAC-seq workflow

1. Global QC

ATAC-seq workflow

1. Global QC

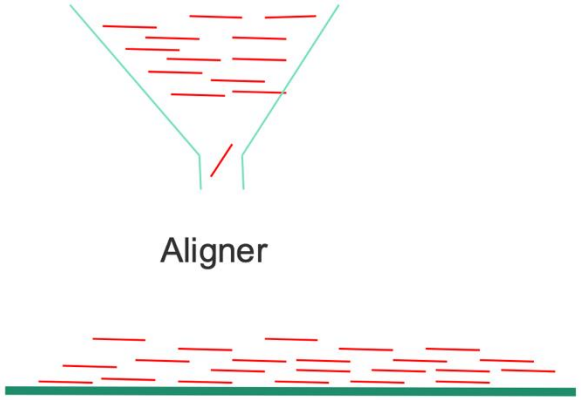
ATAC-seq workflow



1. Global QC
(trimming, removing adapters)

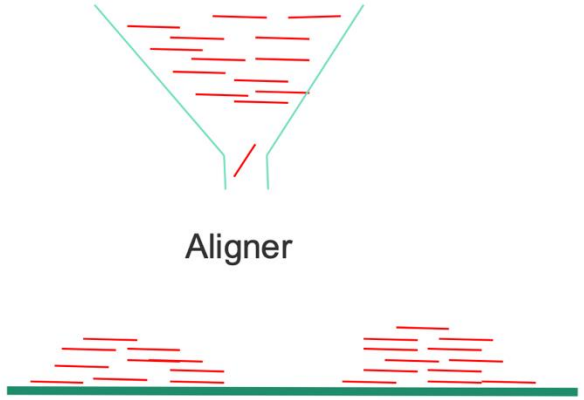
1. Alignment: bowtie, bwa,...

Whole genome re-sequencing
Prokaryote RNAseq



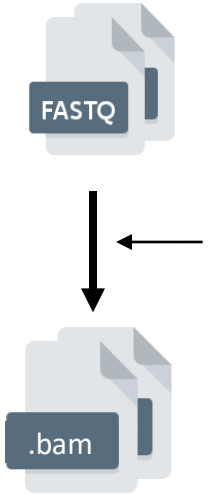
- BWA (Li and Durbin 2009)
- Bowtie (Langemead et al. 2009)

Transcriptome sequencing (RNA-seq)



- Tophat (Trapenell et al. 2009, Kim et al. 2013)
- STAR (Dobin et al. 2013)

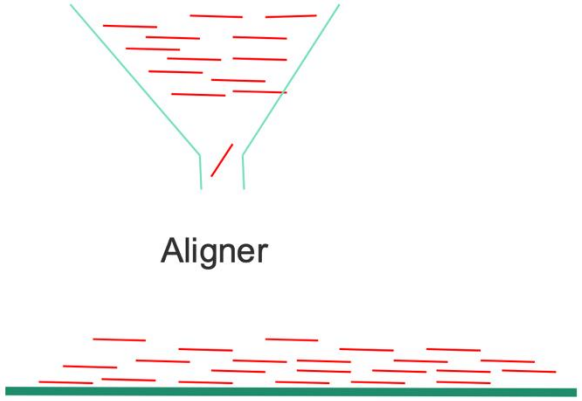
ATAC-seq workflow



1. Global QC
(trimming, removing adapters)

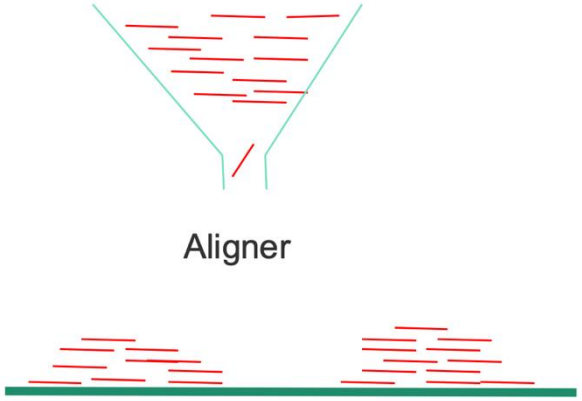
1. Alignment: bowtie, bwa,...

Whole genome re-sequencing
Prokaryote RNAseq



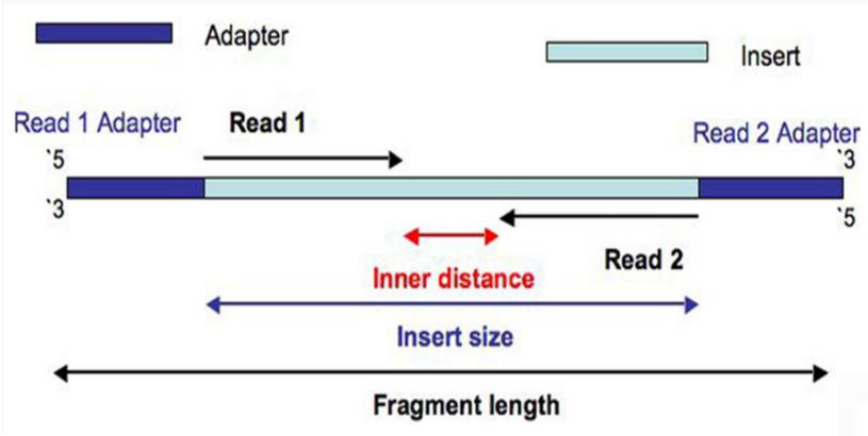
- BWA (Li and Durbin 2009)
- Bowtie (Langemead et al. 2009)

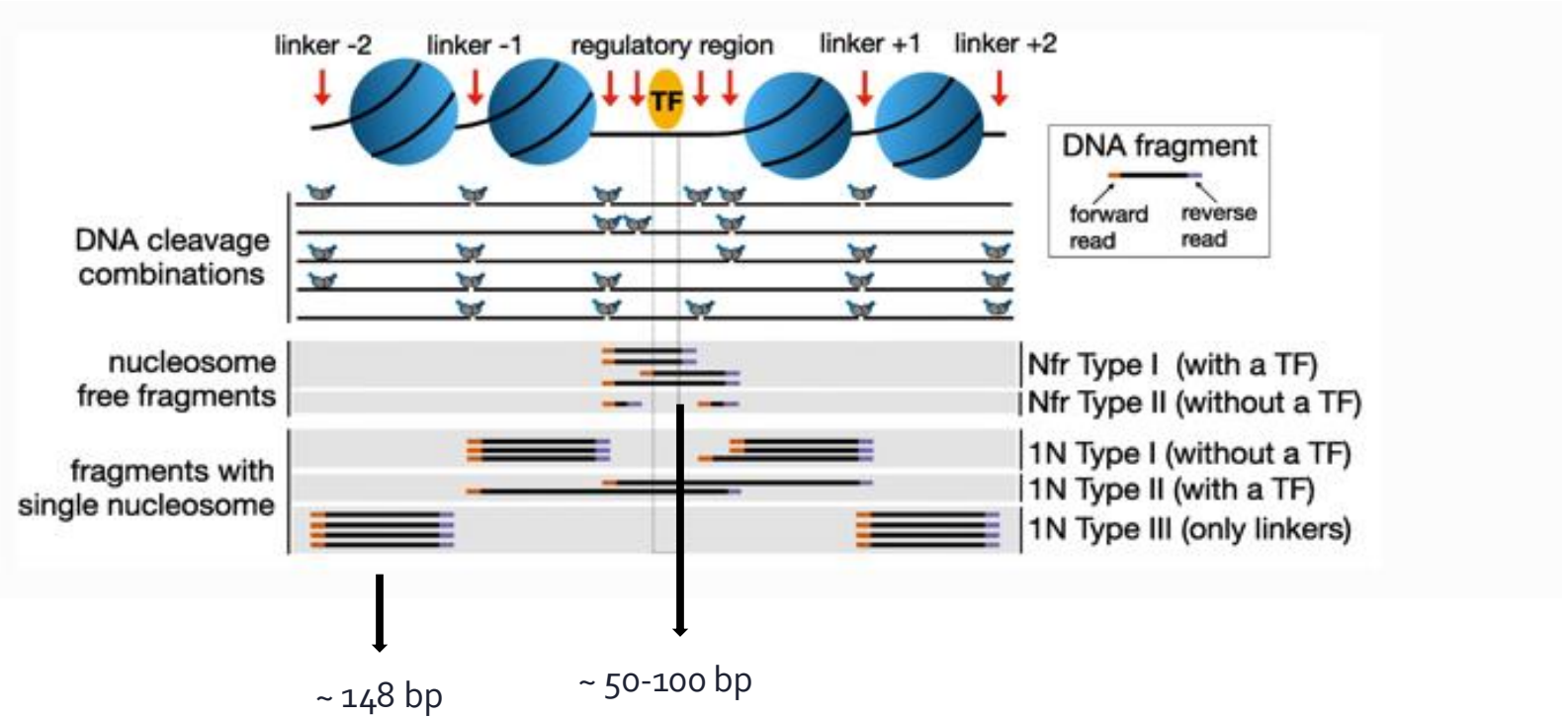
Transcriptome sequencing (RNA-seq)



- Tophat (Trapenell et al. 2009, Kim et al. 2013)
- STAR (Dobin et al. 2013)

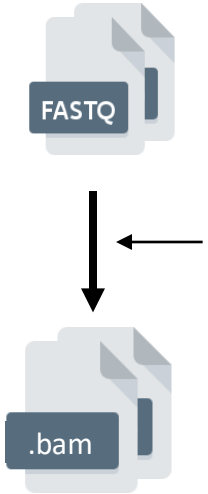
Fragment length / Insert size





Li et al. *Genome Biology*, 2019

ATAC-seq workflow



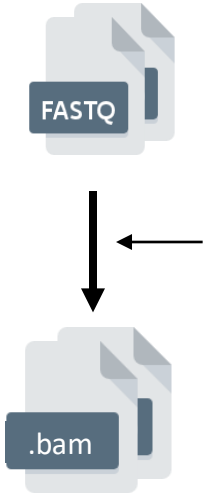
1. Global QC
(trimming, removing adapters)

1. Alignment: bowtie, bwa,…

- **Maximum fragment length** for valid paired-end alignments
Default is 500, could be extended to capture poly-nucleosome fragments

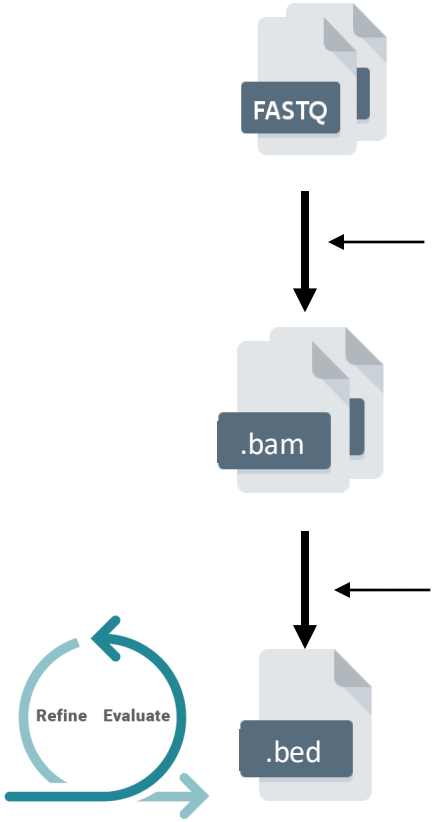
- **End-to-end mode**
We have trimmed the adapters

ATAC-seq workflow

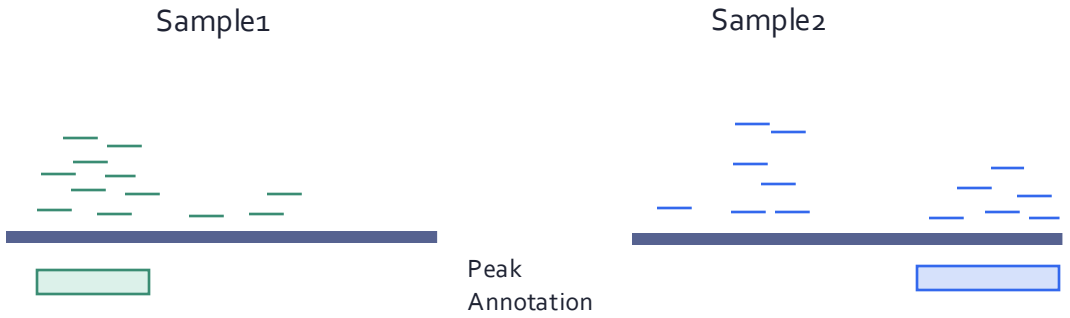


1. Global QC
2. Alignment
3. ATACseq specific QC:
    - Transcription Start Site (TSS) Enrichment Score
    - Fragment length distribution

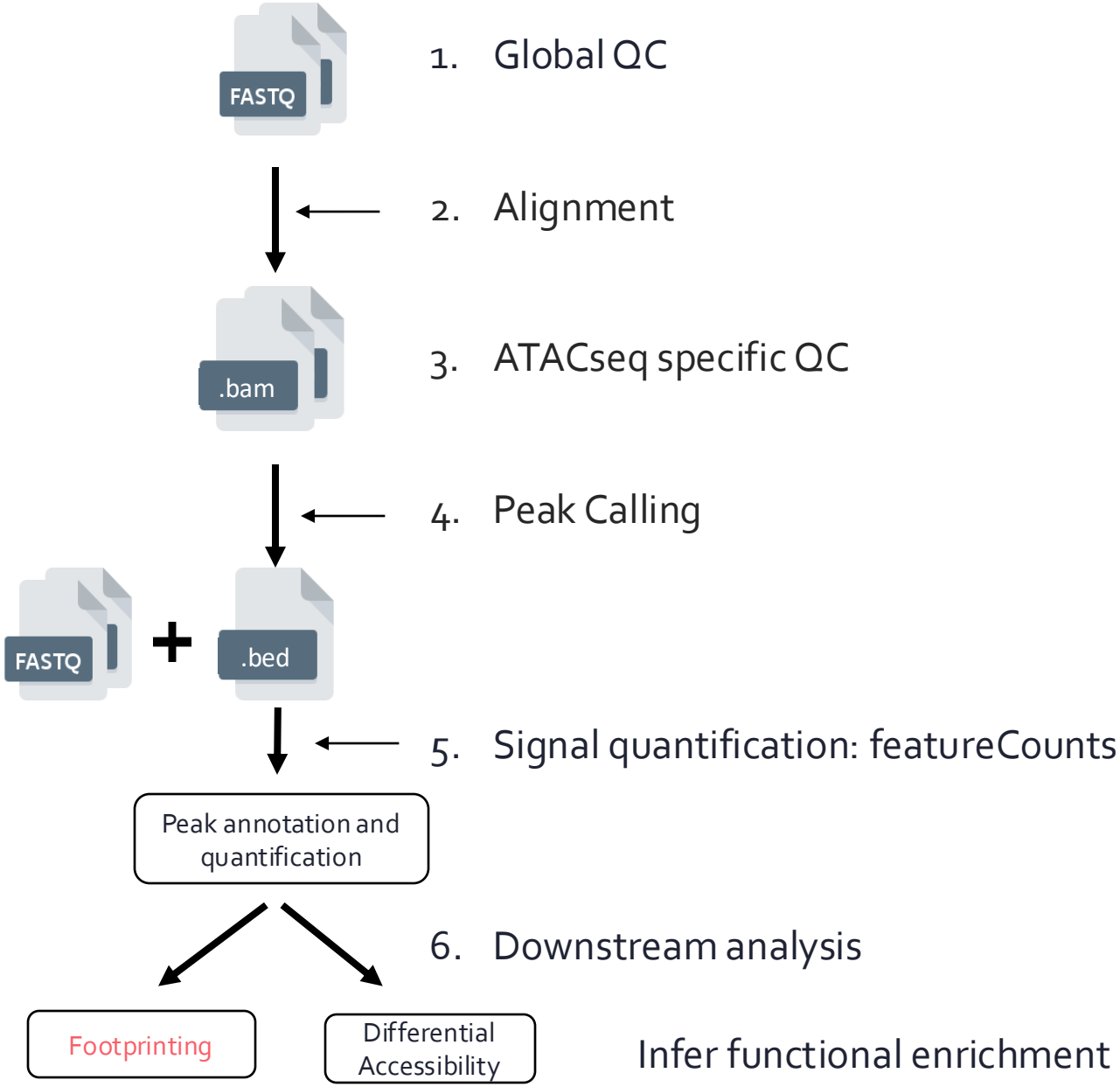
ATAC-seq workflow

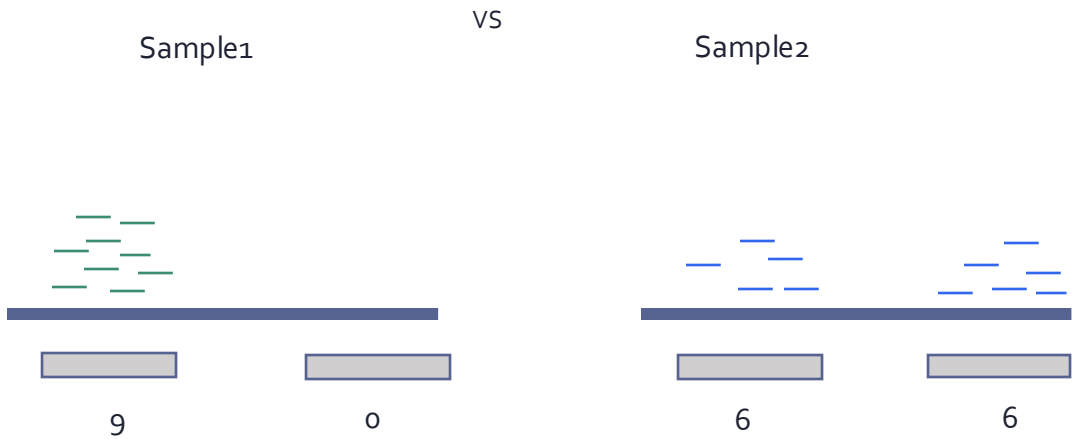


1. Global QC
2. Alignment
3. ATACseq specific QC
4. Peak Calling: macs2, homer, Hmmratac…
   - Adjust parameters (broad vs narrow, pval,…)
   - Define consensus (min. rep.)



ATAC-seq workflow





ATAC-seq experiments

Experimental design

- Cells: 50-500K
- Coverage (50M reads for DA, 200M reads for TF!)
- Controls not needed: Tn5 has some preferred sites that may be confounded, but intrinsic cutting preference is minimal
- Replicates >2 for DA
- Preferentially PE (50-75 bp), SE possible

- Different species and different cell lines may have different behaviour (requirements)

Course exercises

DATASET

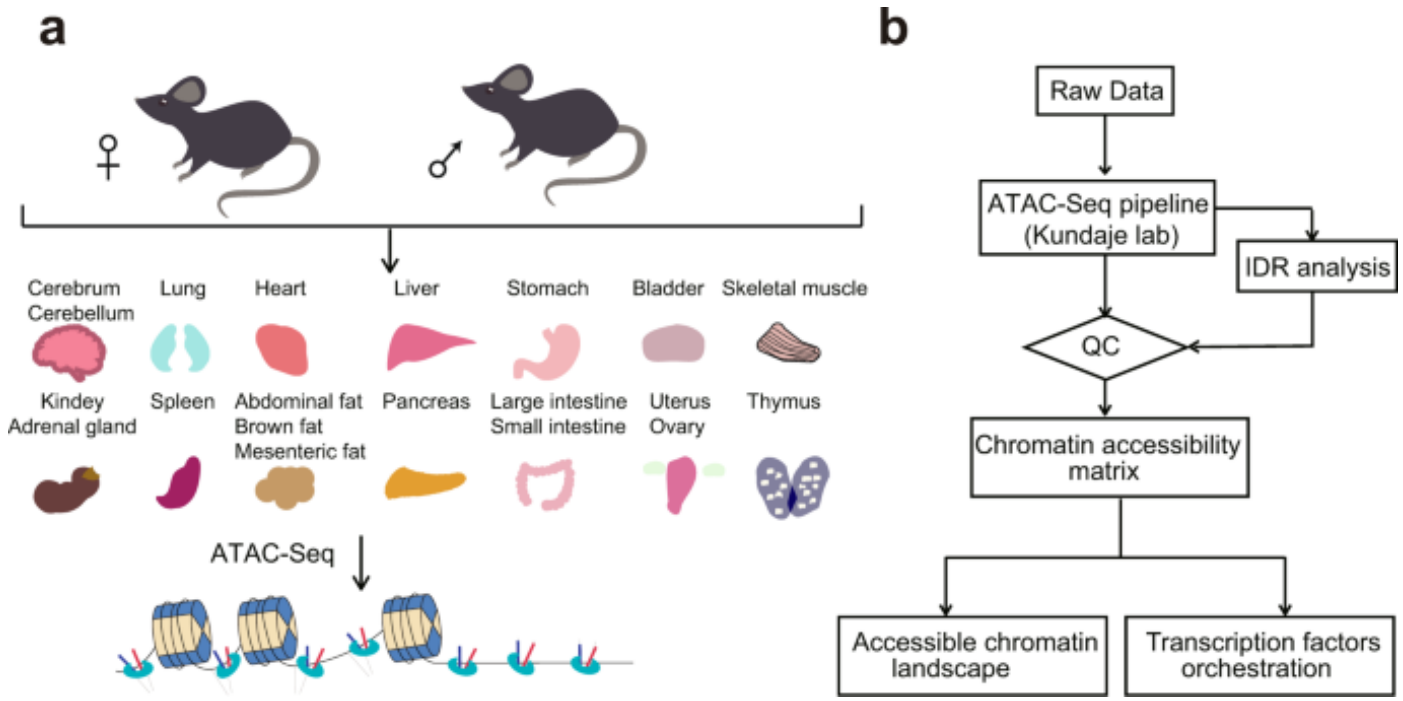
"An ATAC-seq atlas of chromatin accessibility in mouse tissues"

Liu et al. 2019, *Scientidic Data*

Female Kidney (2x)

vs

Female Cerebrum (2x)



Course exercises

DATASET

"An ATAC-seq atlas of chromatin accessibility in mouse tissues"
Liu et al. 2019, *Scientidic Data*

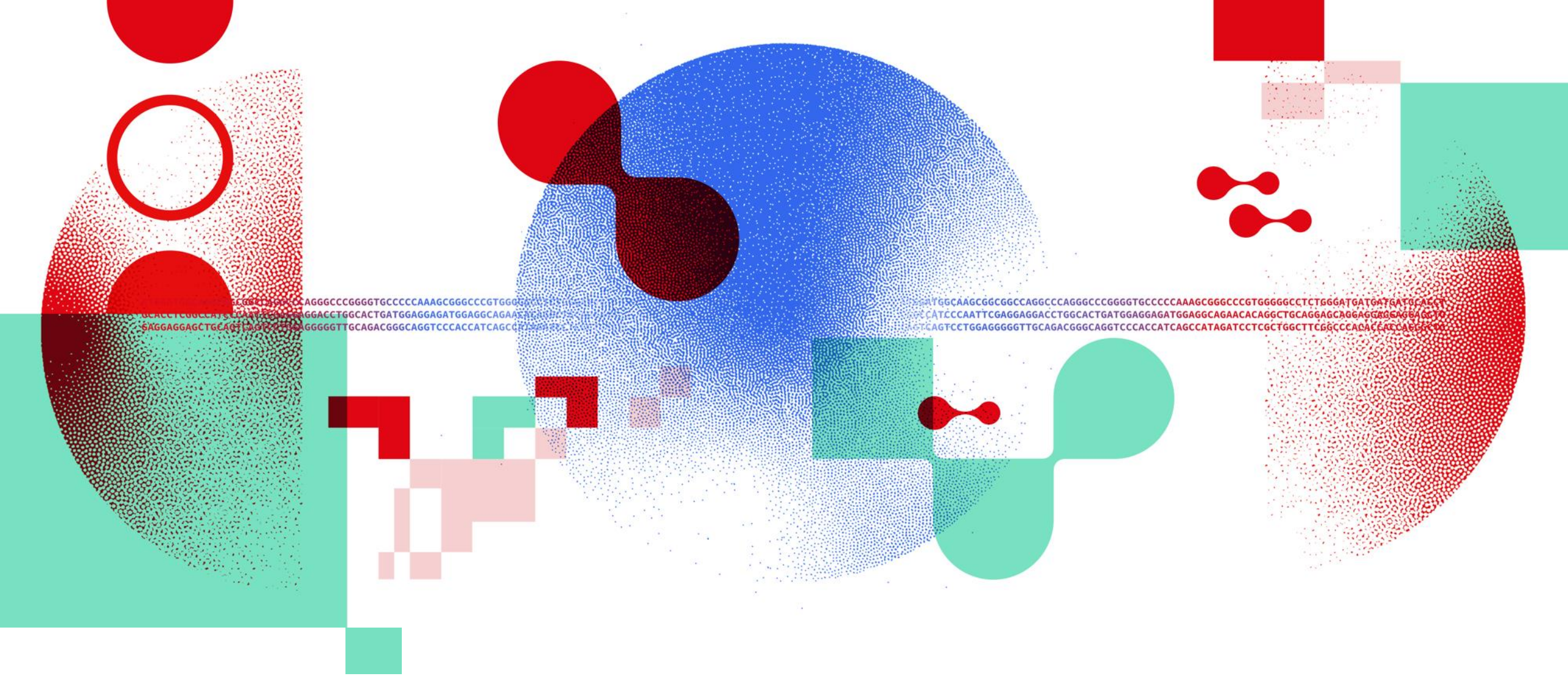
Female Kidney (2x)

vs

Female Cerebrum (2x)

.bam files

QC report



QUESTIONS?

DATA SCIENTISTS FOR LIFE

sib.swiss