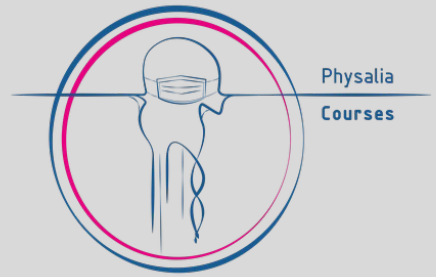
ATAC-seq analysis

Epigenomics Data Analysis

Jacques Serizay

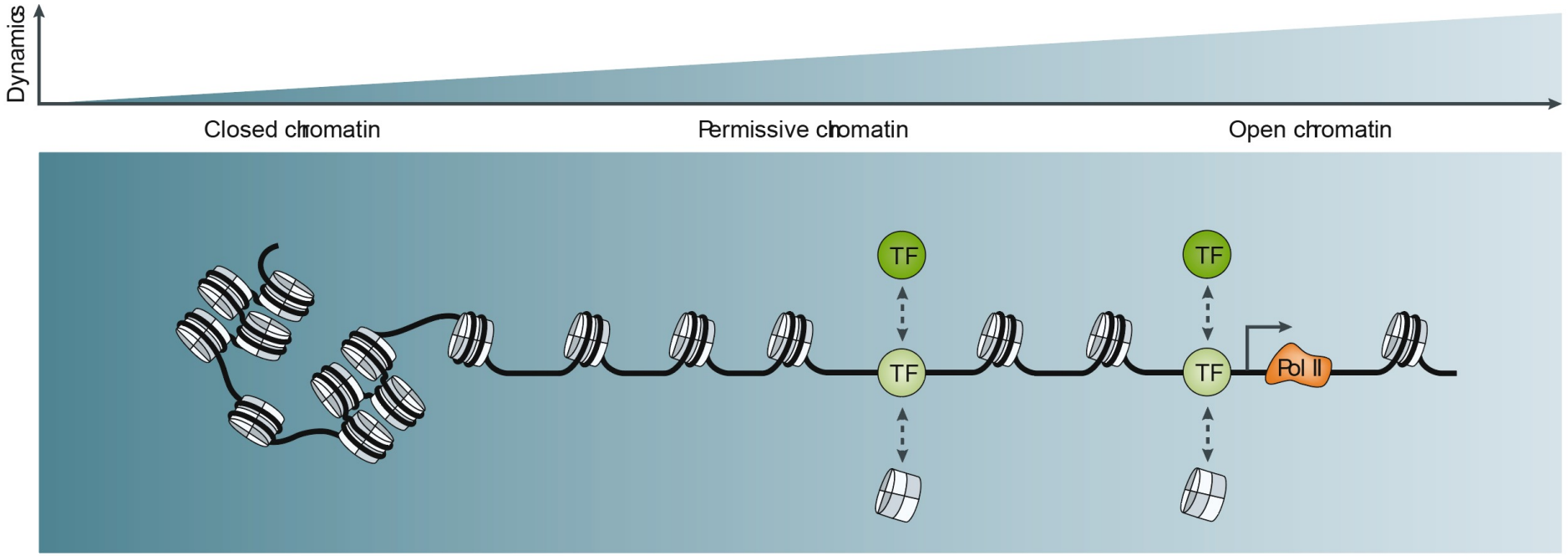
Physalia 2025



Beyond textbook statements: a spectrum of regulatory capacity

Chromatin accessibility continuum that ranges from closed chromatin to highly dynamic, accessible or permissive chromatin

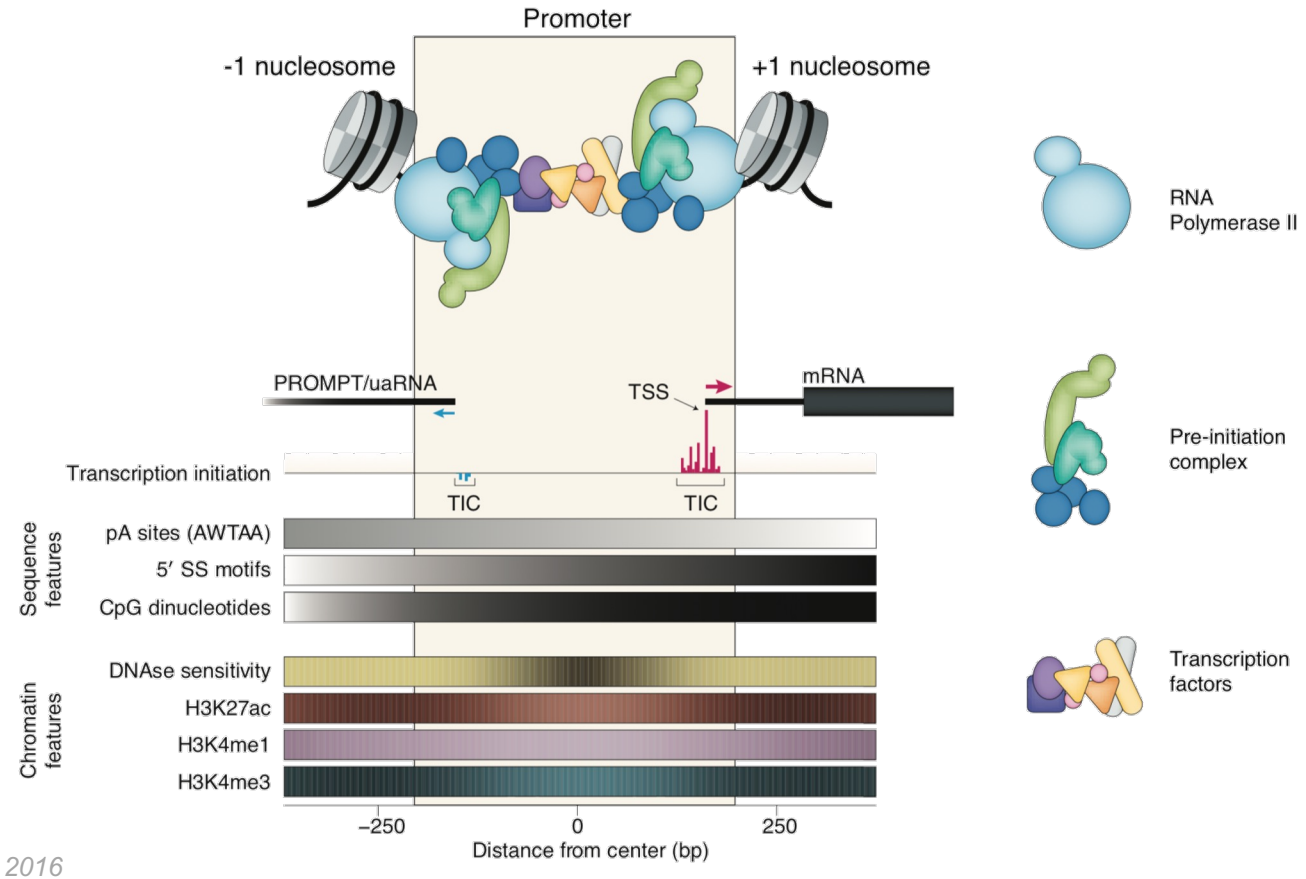
This landscape of chromatin accessibility reflects the spectrum of regulatory capacity — rather than a bistate organization



Klemm et al., Nat .Rev. Genet 2019

Promoter organization

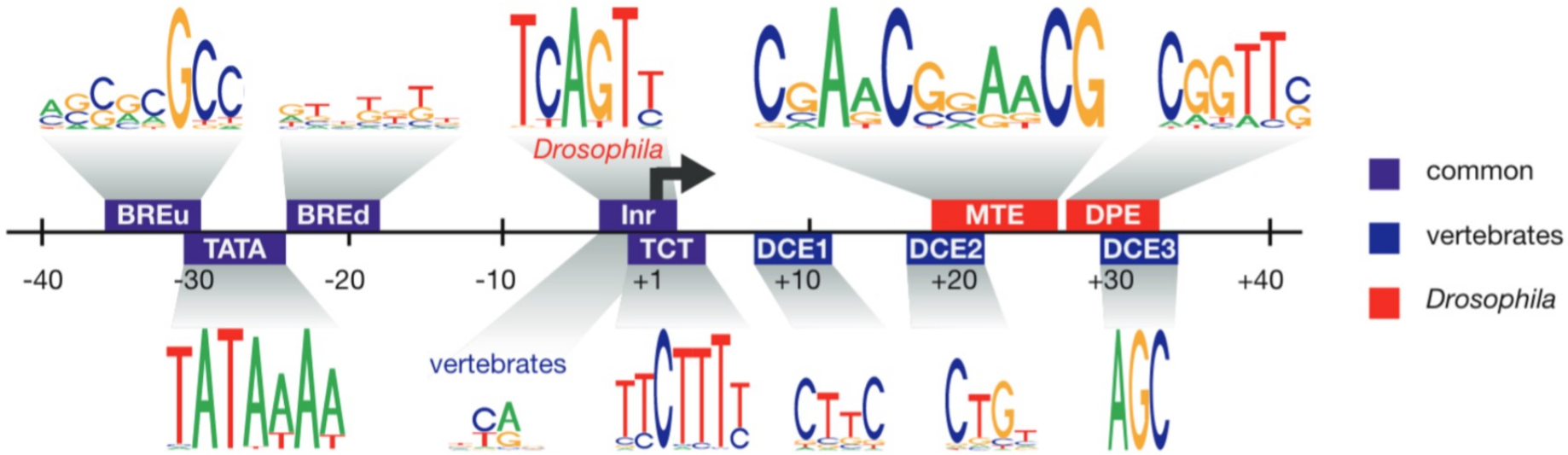
Promoters are crowded environments!



Haberle and Lenhard, Semin. Cell Dev. Biol. 2016

Promoter organization

Transcription machinery and general transcription factors need access to DNA to recognize their binding motif

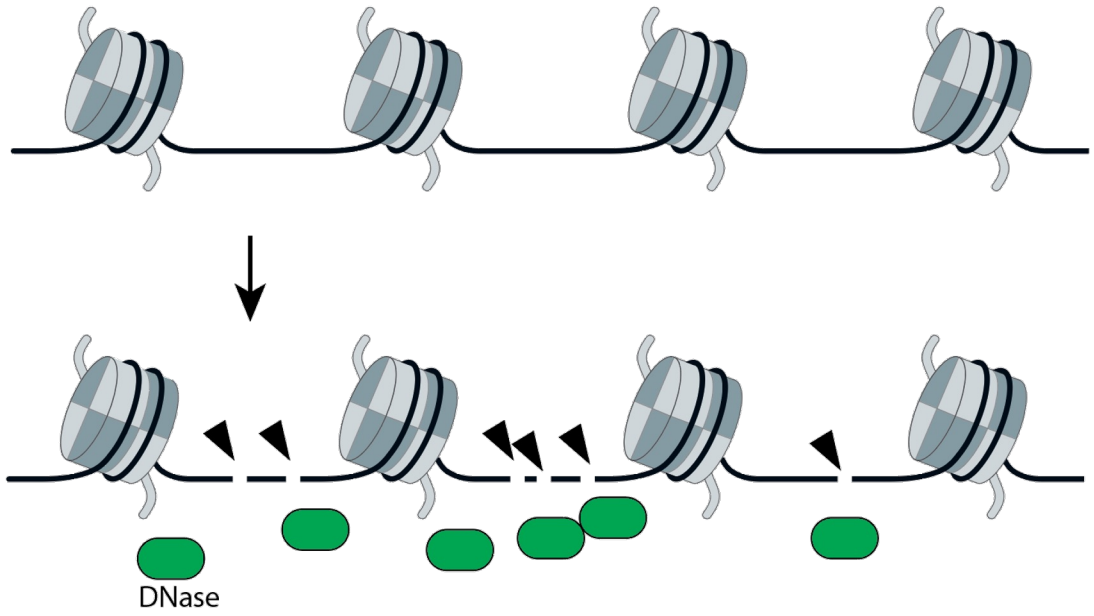


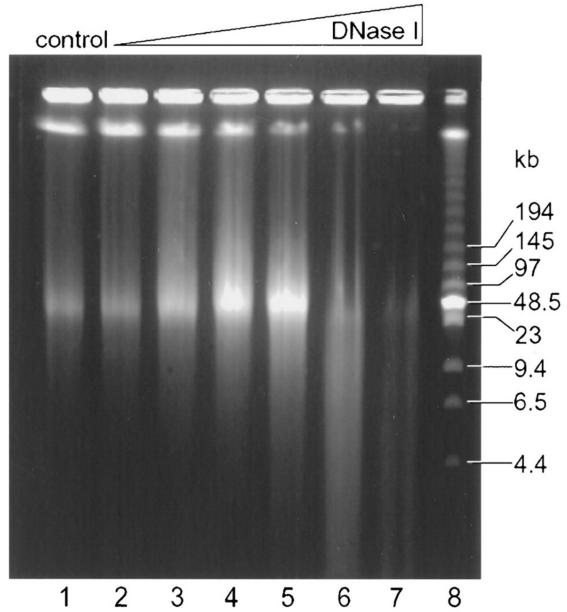
Haberle and Lenhard, Semin. Cell Dev. Biol. 2016

How to measure chromatin accessibility: originally with nucleases

Nuclease enzymes were historically used to profile chromatin accessibility

- Dnase I (deoxyribonuclease I): a non-specific double-strand endonuclease



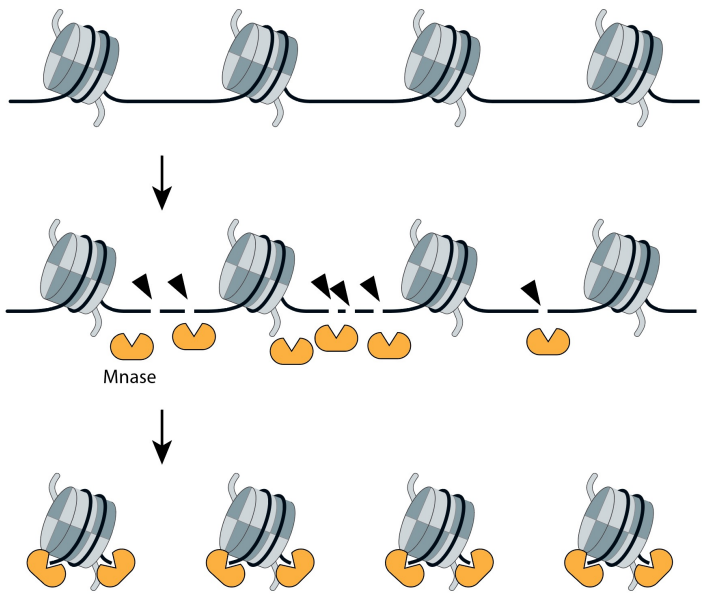


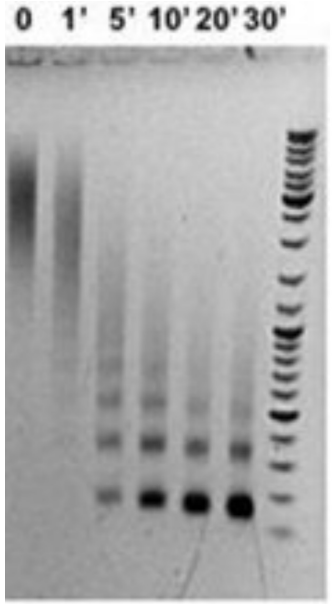
Paul & Ferl, Plant Cell 1998

How to measure chromatin accessibility: originally with nucleases

Nuclease enzymes were historically used to profile chromatin accessibility

- Dnase I (deoxyribonuclease I): a non-specific double-strand endonuclease
- Mnase (Micrococcal nuclease): a endo-exonuclease with preference to single-strand DNA but with also double-strand nuclease activity



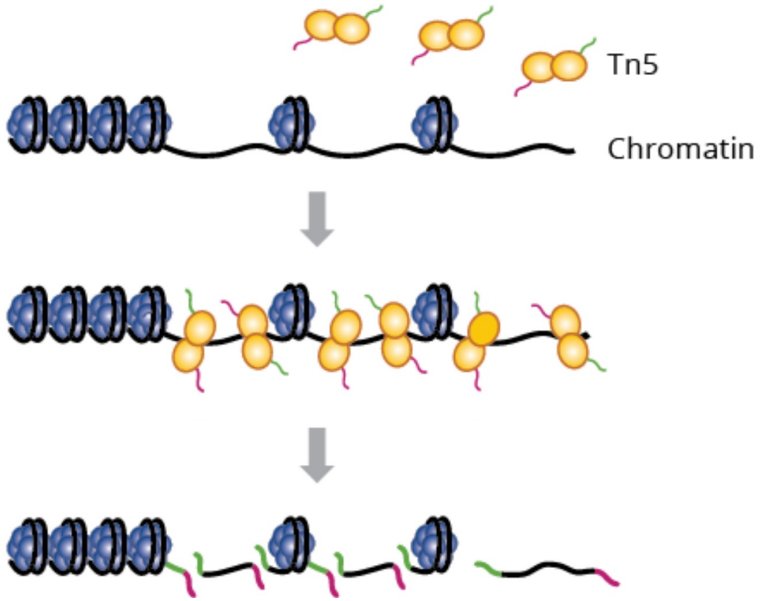


Rodríguez-Campos & Azorín, PLoS One 2007

How to measure chromatin accessibility: originally with nucleases

Nuclease enzymes were historically used to profile chromatin accessibility

- Dnase I (deoxyribonuclease I): a non-specific double-strand endonuclease
- Mnase (Micrococcal nuclease): a endo-exonuclease with preference to single-strand DNA but with also double-strand nuclease activity
- ATAC: a Tn5 transposase, integrating transposons wherever it is possible (i.e. accessible)



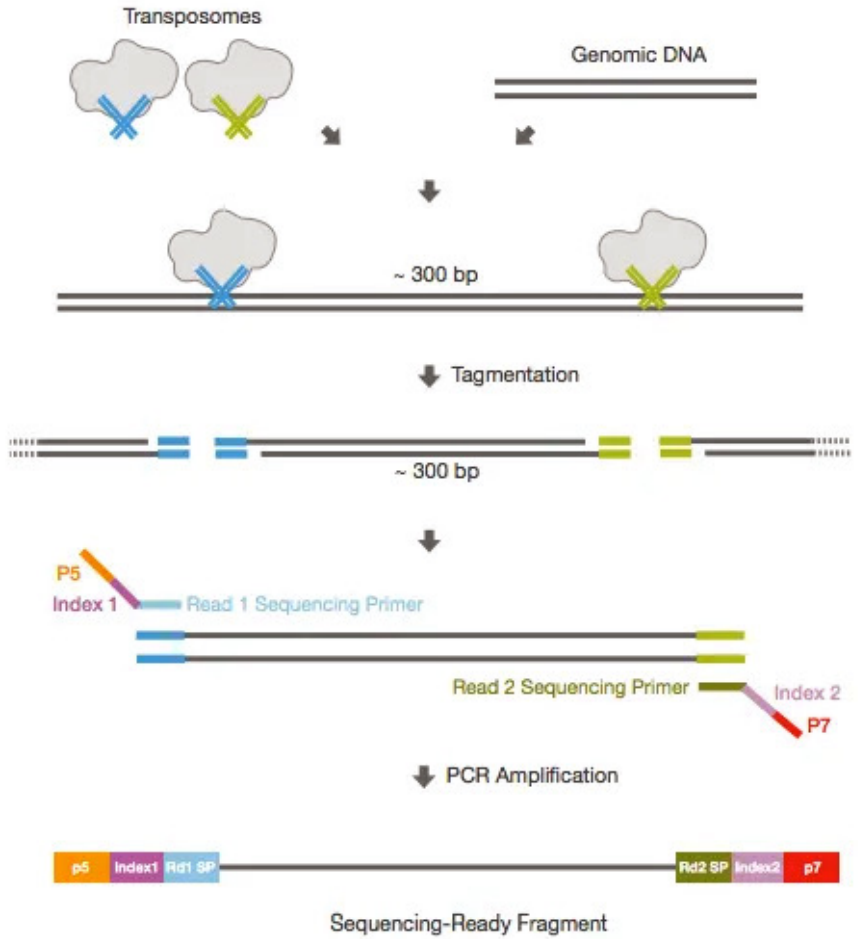
Buenrostro et al., 2013

ATAC-seq: from chromatin to NGS library

Since the sequence of the transposons loaded on the Tn5 transposome is known, one can use them to in a PCR

→ “Tagmented” (i.e. DNA with inserted transposons) will be amplified.

→ Each end of a fragment corresponds to a transposition event.

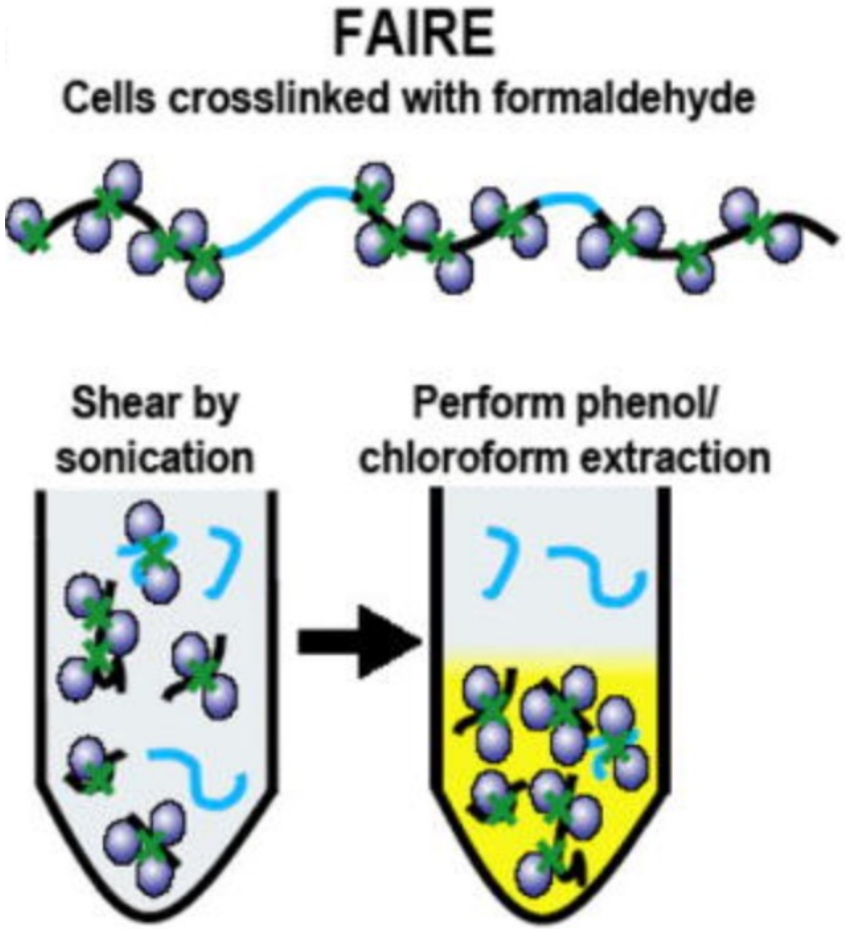


Emergence of other enzymatic or mechanical approaches

Mechanical approaches

- **FAIRE-seq**: Formaldehyde-Assisted Isolation of Regulatory Elements

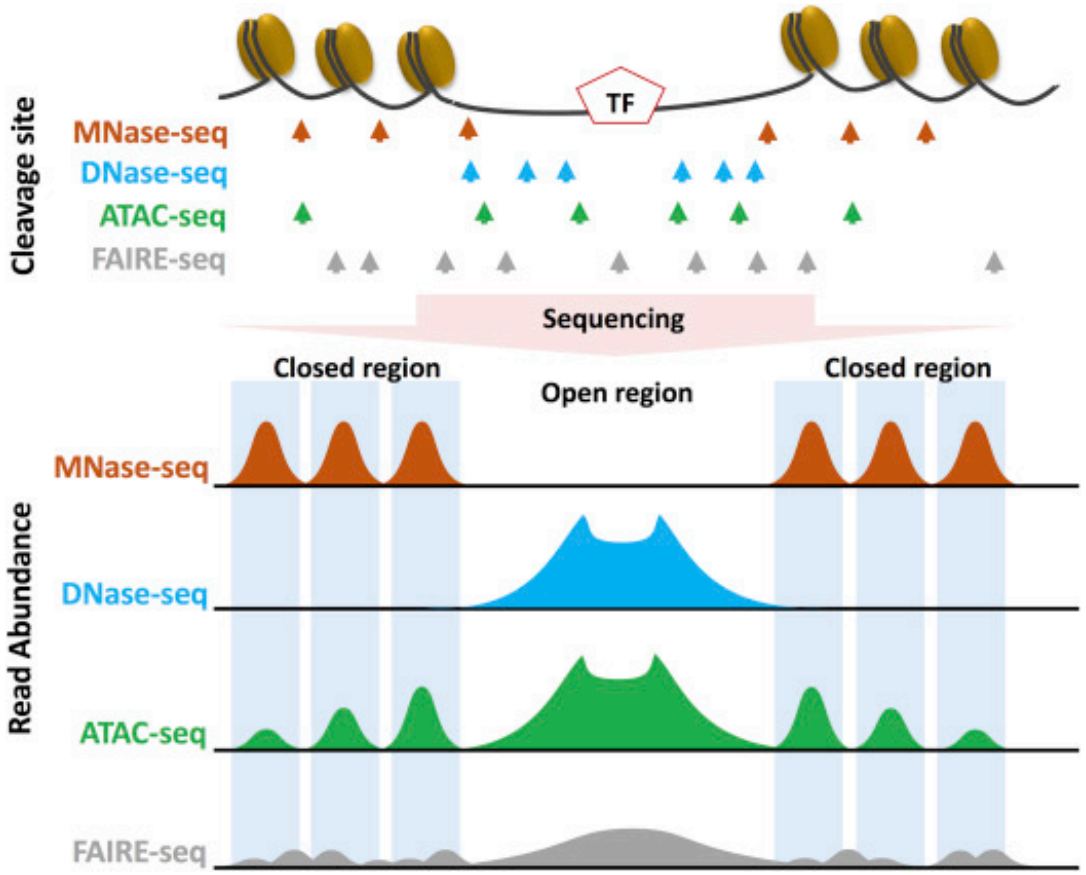
→ Uses a crosslinking + sonication + phenol extraction to isolate nucleosome-depleted chromatin



Giresi & Lieb, Methods 2009

Comparison of the main experimental approaches

Each assay generates a specific type of profile.

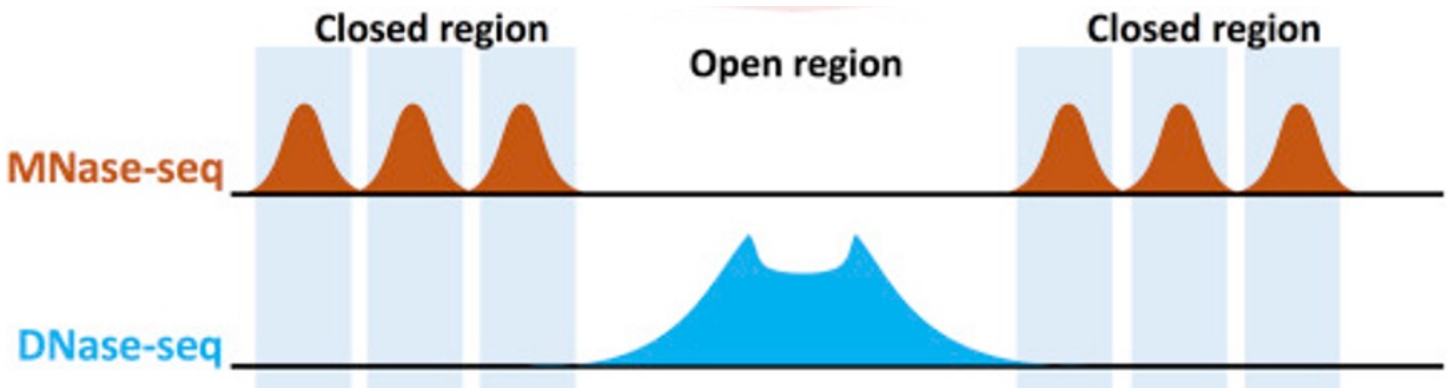


Hsu et al., Epigenetics in Human Diseases 2018

Positive vs. negative measurements

DNAse-seq relies on presence of signal (**positive measurements**) to map accessible regulatory elements

Mnase-seq relies on absence of signal (**negative measurements**) to map accessible regulatory elements



ATAC-seq downstream analysis

- ✓ Get .bcl files 
- ✓ Create fastq files  Or **`bcl2fastq`**
- ✓ QC: remove/trim low quality reads — E.g. **`cutadapt`**
- ✓ Align fastq to BAM — E.g. **`bowtie2`**
- ✓ Filter duplicates, artifacts, … — E.g. **`samtools`**
- ✓ Generate tracks — E.g. **`deepTools`**
- ● Assay-specific downstream analysis

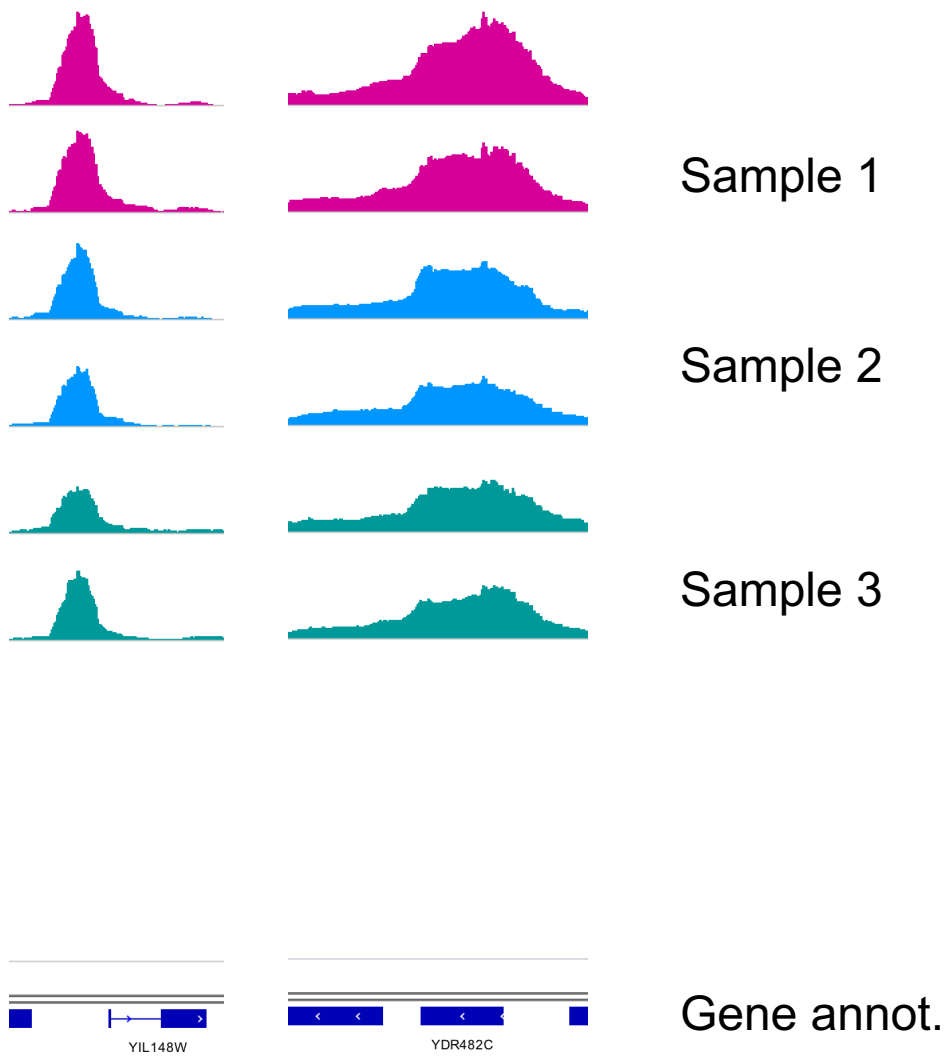
Peak calling
Differential peak analysis
Fragment size distribution
Promoter footprint analysis

ATAC peak callers

- Most peak callers were designed before the emergence of ATAC-seq
- Few of them directly aim at identifying peaks in chromatin accessibility signals

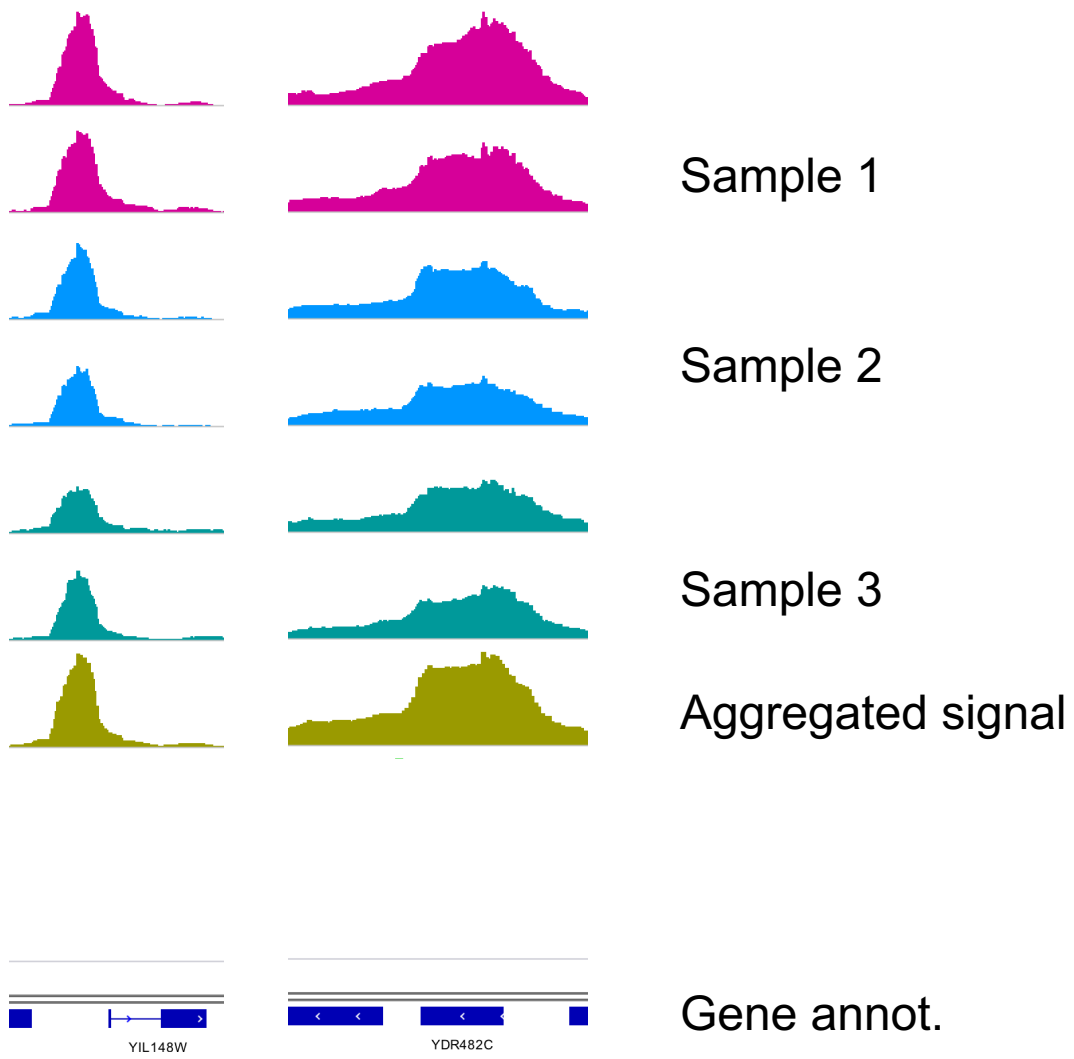
ATAC peak callers

- Most peak callers were designed before the emergence of ATAC-seq
- Few of them directly aim at identifying peaks in chromatin accessibility signals
- Since the emergence of ATAC-seq, new peak callers were designed with the identification of peak “shape” in mind.



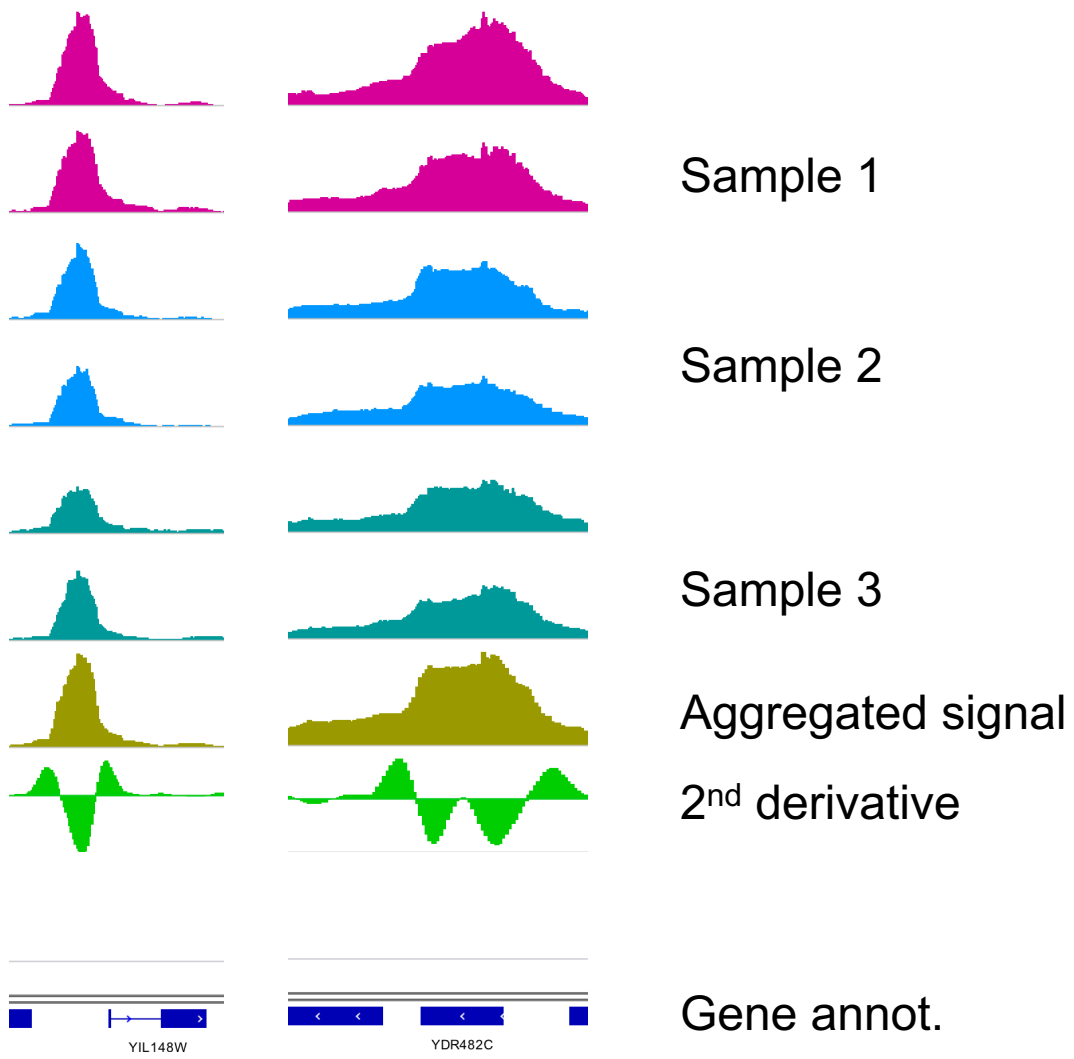
ATAC peak callers

- Most peak callers were designed before the emergence of ATAC-seq
- Few of them directly aim at identifying peaks in chromatin accessibility signals
- Since the emergence of ATAC-seq, new peak callers were designed with the identification of peak “shape” in mind.



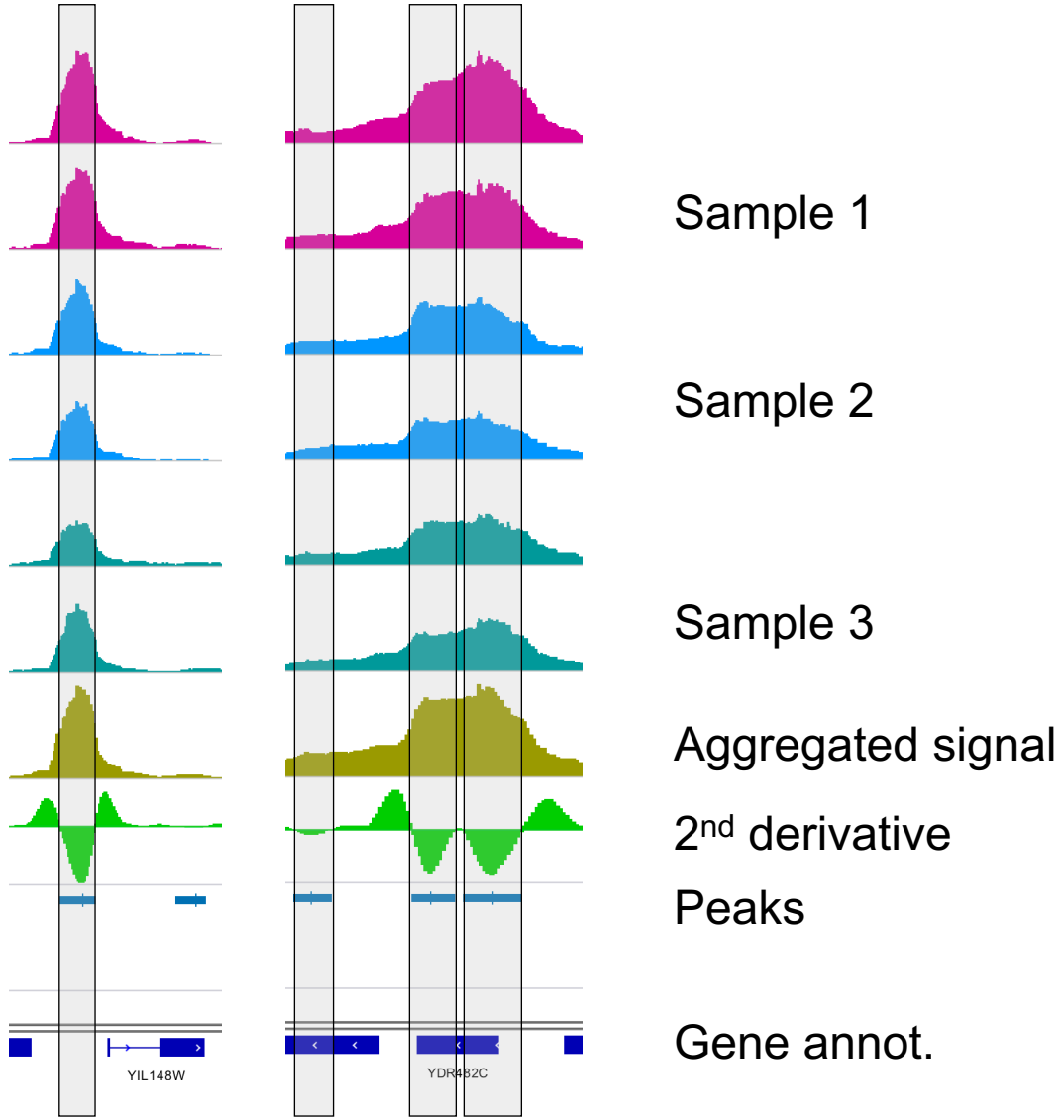
ATAC peak callers

- Most peak callers were designed before the emergence of ATAC-seq
- Few of them directly aim at identifying peaks in chromatin accessibility signals
- Since the emergence of ATAC-seq, new peak callers were designed with the identification of peak “shape” in mind.



ATAC peak callers

- Most peak callers were designed before the emergence of ATAC-seq
- Few of them directly aim at identifying peaks in chromatin accessibility signals
- Since the emergence of ATAC-seq, new peak callers were designed with the identification of peak “shape” in mind.



ATAC peak callers

- Most peak callers were designed before the emergence of ATAC-seq
- Few of them directly aim at identifying peaks in chromatin accessibility signals
- Since the emergence of ATAC-seq, new peak callers were designed with the identification of peak “shape” in mind.

```
yapc output_refix \
  sample1 sample1_rep1.bw sample1_rep2.bw \
  sample2 sample2_rep1.bw sample2_rep2.bw \
  … \
  sampleN sampleN_rep1.bw sampleN_rep2.bw
```