Chromotin modification, ChIP-seq and DNA accessibility sequencing methods

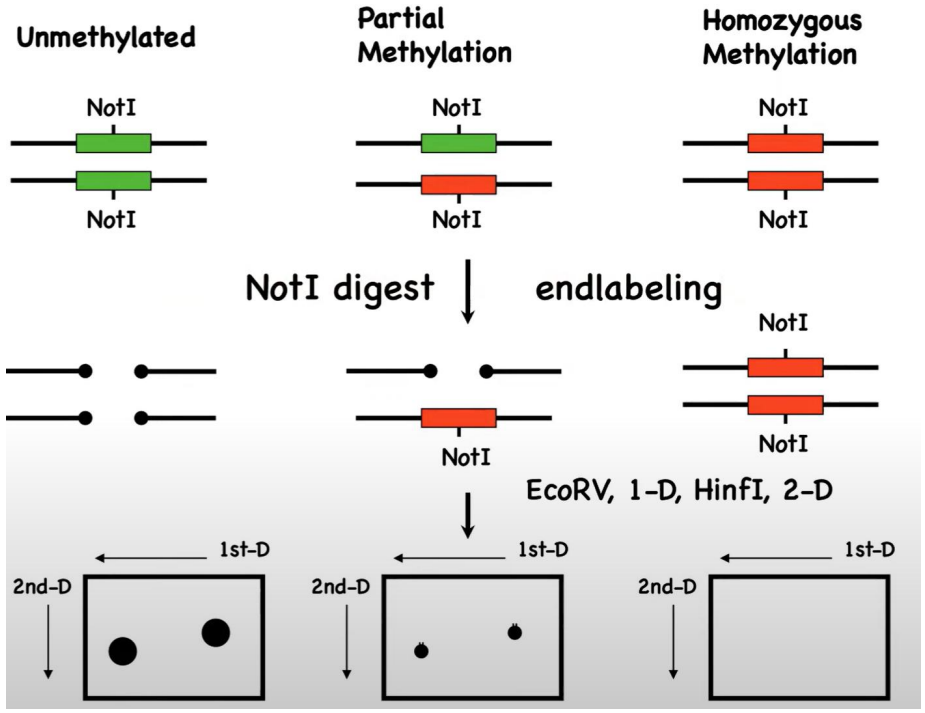
Qiayi zha 3.12

DNA Methylation Detection

part1

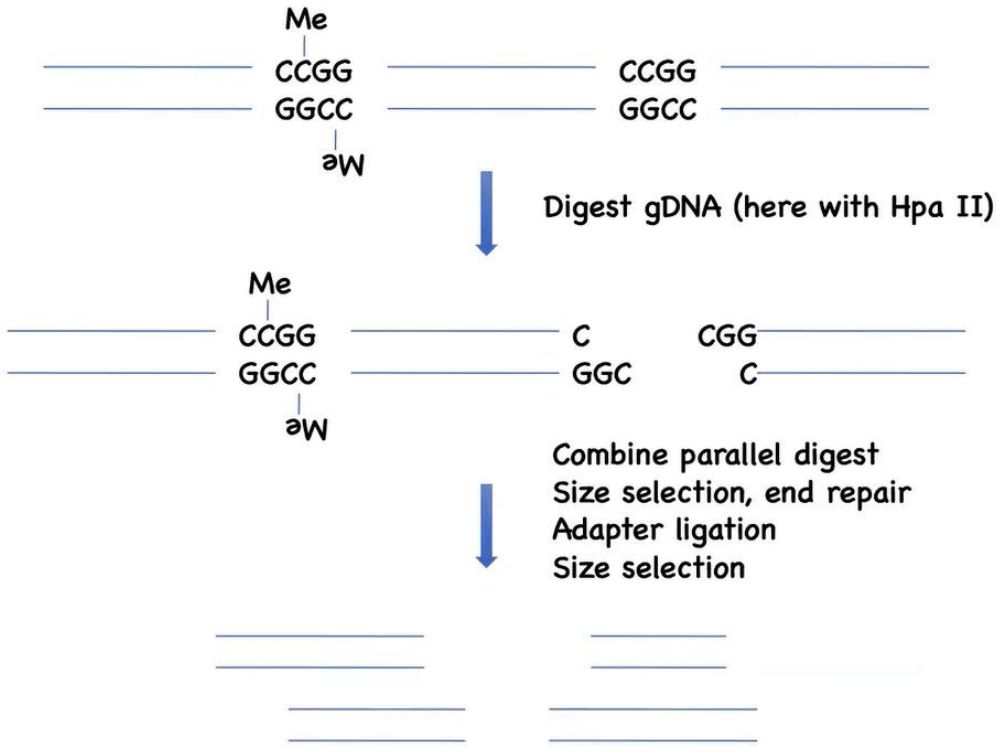
1.1 Detect/measure DNA methylation: Restriction enzymes

- Restriction landmark genomic scanning (RLGS)
- Use restriction enzymes some of which are specific to DNA modifications



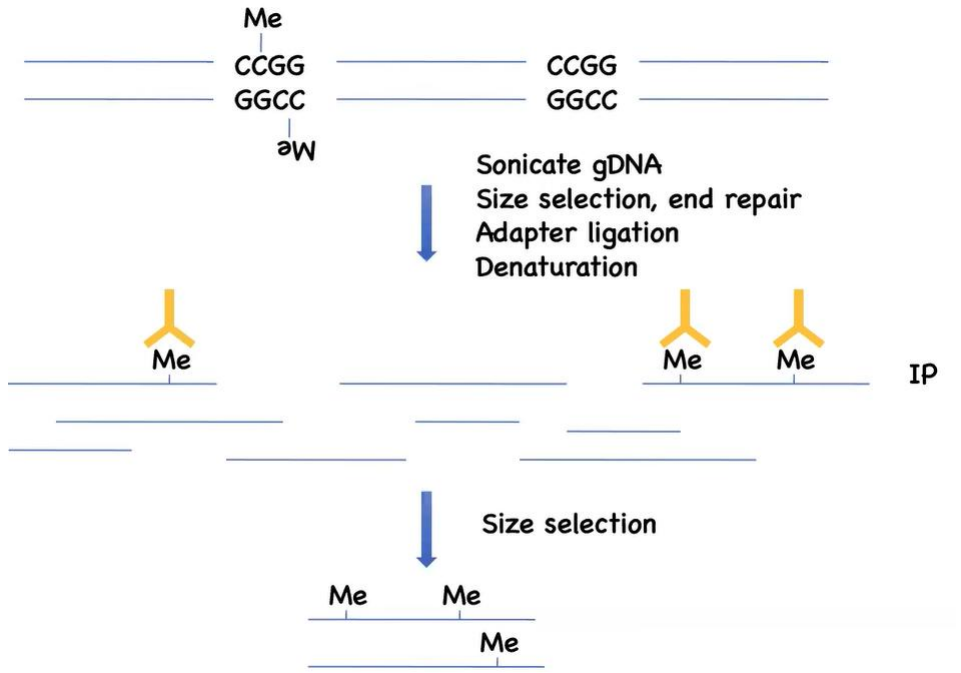
1.2 Detect/measure DNA methylation

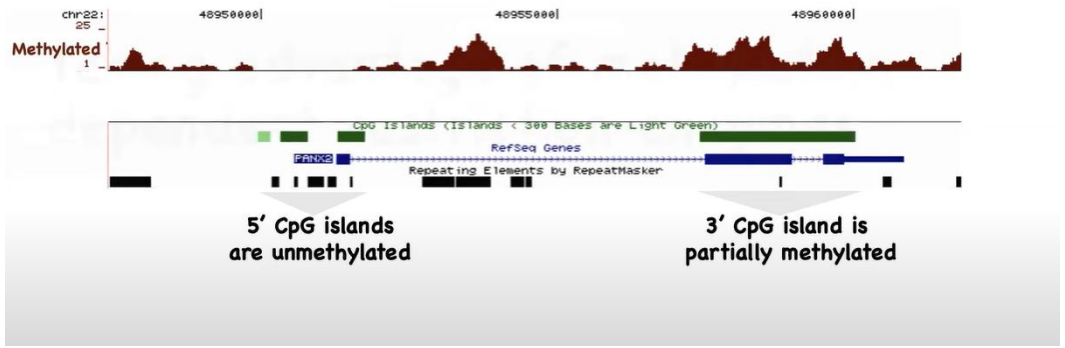
- Methylation sensitive restriction enzyme sequencing (MSRE/MRE-Seq)
- Use methylation sensitive restriction enzymes on genomic DNA to study DNA methylation



1.3 Detect/measure DNA methylation: Immunoprecipitation

- Methylated RNA immunoprecipitation sequencing (MeRIP-seq, m6A-seq)
- Detection of post-transcriptional RNA modifications

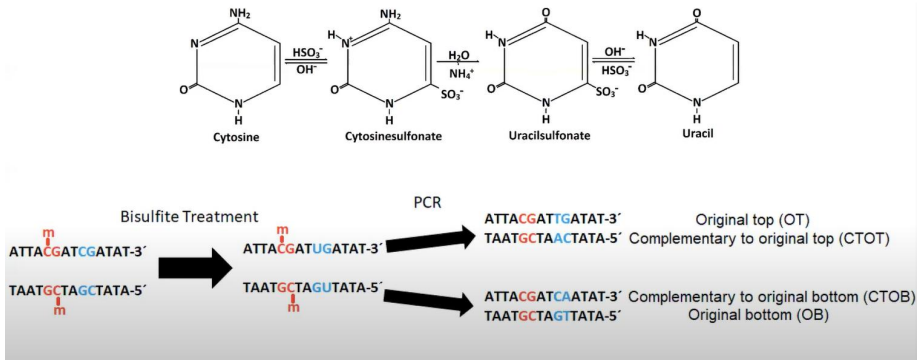


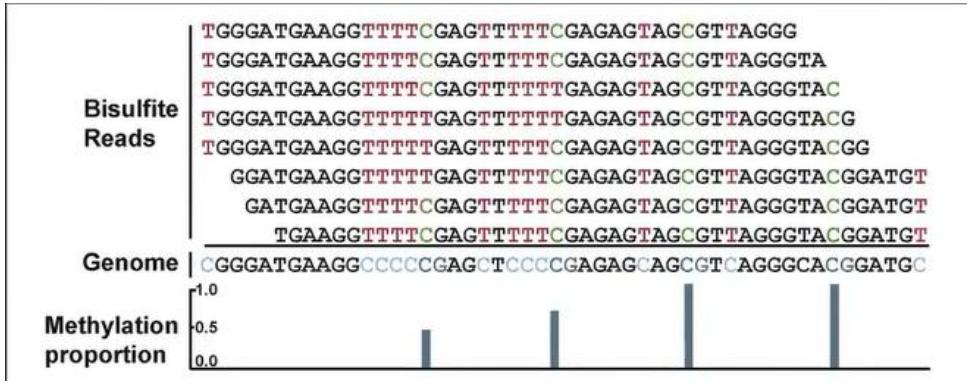


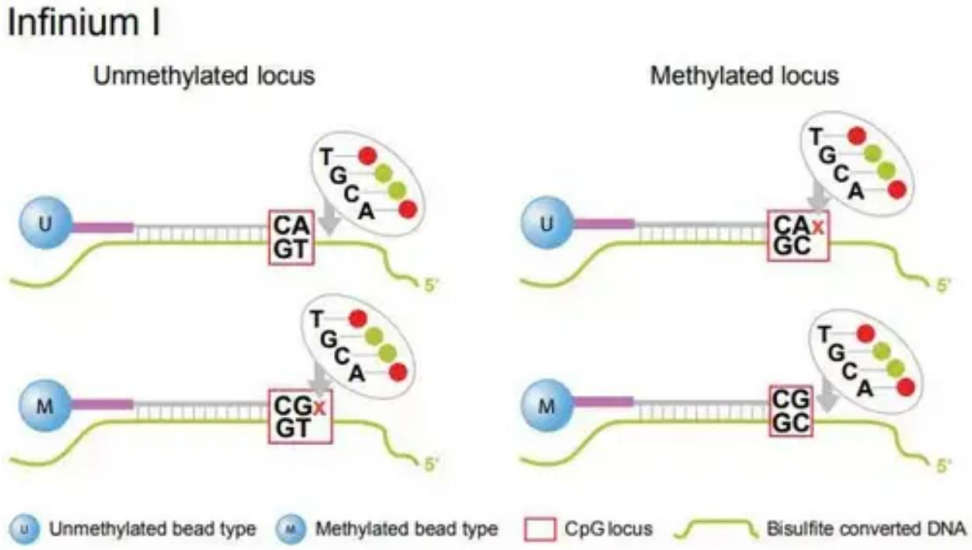
Li et al, Methods, 2014

1.4 Detect/measure DNA methylation: 5m-C

- Bisulfite sequencing (BS-Seq) / Whole-genome bisulfite sequencing (WGBS)
- Detect methylated cytosines in genomic DNA -> illumina Beadchip



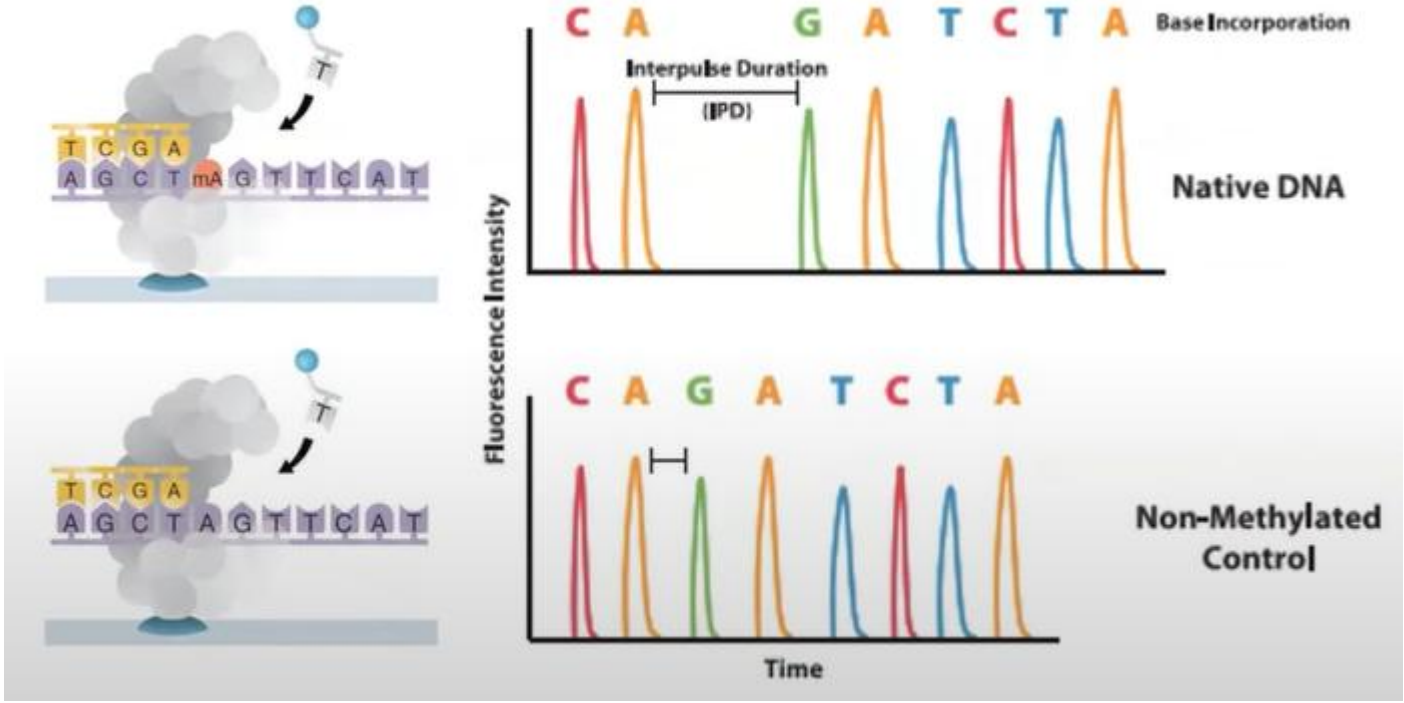




Grehl et al., Epigenetics, 2018

1.5 Detect/measure DNA methylation: polymerase kinetics

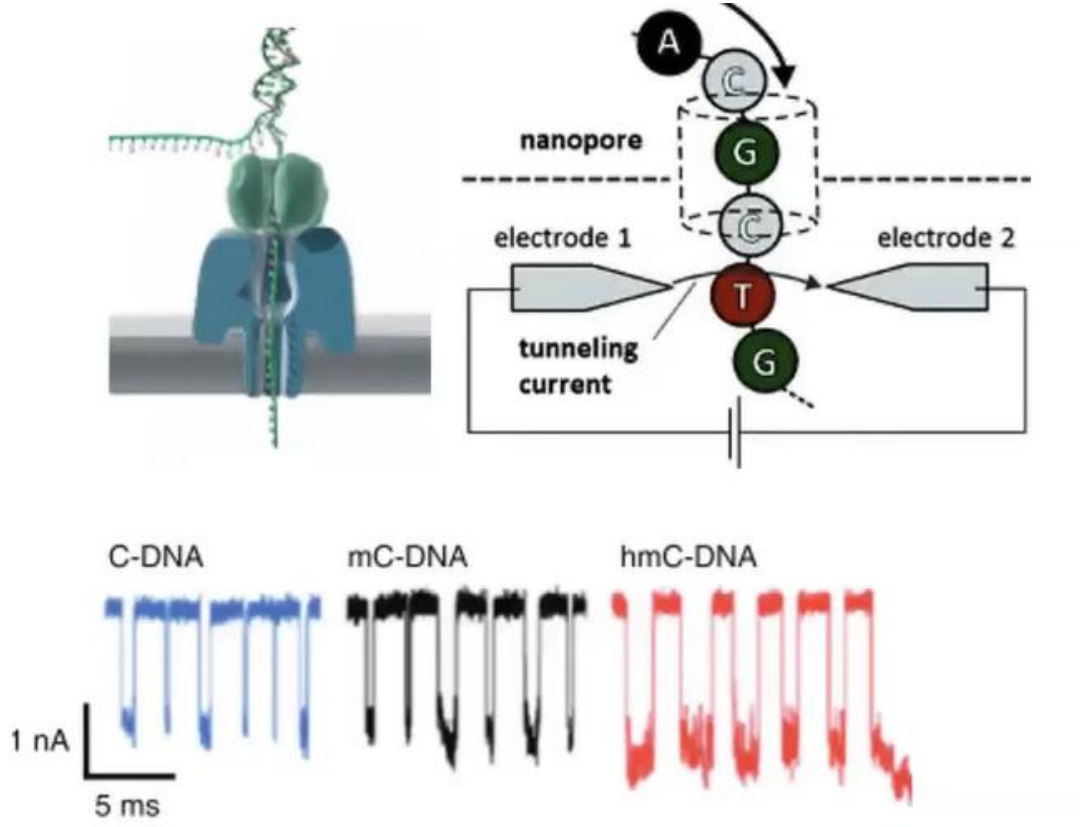
- Single-molecule real-time (SMRT) sequencing
- Discriminate between different bases by analyzing variations in polymerase kinetics



Kinney et al., Epigenetic Alterations in Oncogenesis (2013)

1.6 Detect/measure DNA methylation: Nanopore

- Nanopore sequencing
- Scan discriminate between C, 5-mC and 5-hmC due to differences incurrent profiles



Korlach & Tuner Current Opinion in Strucfural Biology (2012)

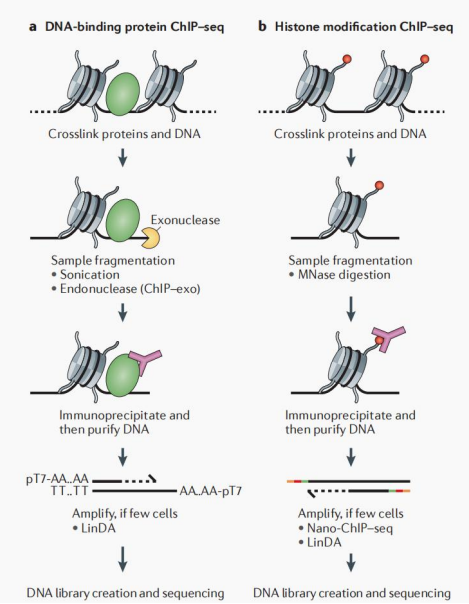
Histone modification examples

- Gene body mark: H3K36me3 (Pol2 elongation), H3K79me3 (first exon / intron)
- Promoter (TSS) mark: H3K4me3
- Enhancer (TF binding) mark: H3K4me1
- Both enhancers and promoters: H3K4me2, H3/H4ac (esp H3K27ac), H2AZ
- Repressive mark: H3K27me3 (CpG islands near promoters), H3K9me3 (repeat region in the genome) → longer term suppression by DNA methylation
- Overall resolution lower than TF ChIP-seq

Protein–DNA binding site identification

part2

2.1 ChIP-Seq



- Chromatin immunoprecipitation sequencing (ChIP-Seq)
- **Histone modification (marks) -> Chromotin regulation**
- 200-600bp fragment

Pros: antibody --- protein or DNA sequence

- protein, TF (antibody - protein or DNA sequence), histone --- difference

Cons:

- Requires **a large amount of input material**, cells or tissue, to produce a strong enough signal over background noise.

- **Cross-linking:** The reliance on fixation increases the chance of false-positives: proteins may bind at particular genomic sites when in fact they are just non-specific or indirectly interacting.

- **Sonication & MNase:** Uneven cut & prefer to cut base A/T

Landt et al., 2012

2.2 Cut&run

- Cleavage Under Targets and Release Using Nuclease(CUT&RUN);
- in situ mapping of protein-DNA interactions
- **pAG-MNase**
- TF fragment sizes peaked: ~100 bp; H2A fragments peaked: ~150 bp

Pros:

- Fast, **require small cell number** (as few as 100-1000 cell)
- High resolution and ~1/10th the sequencing depth as ChIP
- **Yield precise TF profiles**
- **Avoid crosslinking and solubilization issues**

Figure 1: CUT&RUN schematic.

Cons:

- DNA end polishing & adapter ligation: **increases time, cost and efforts**
- Release of MNase-cleaved fragments into the supernatant **not well-suited for single-cell platforms**

Skene, 2017

2.3 Cut&tag

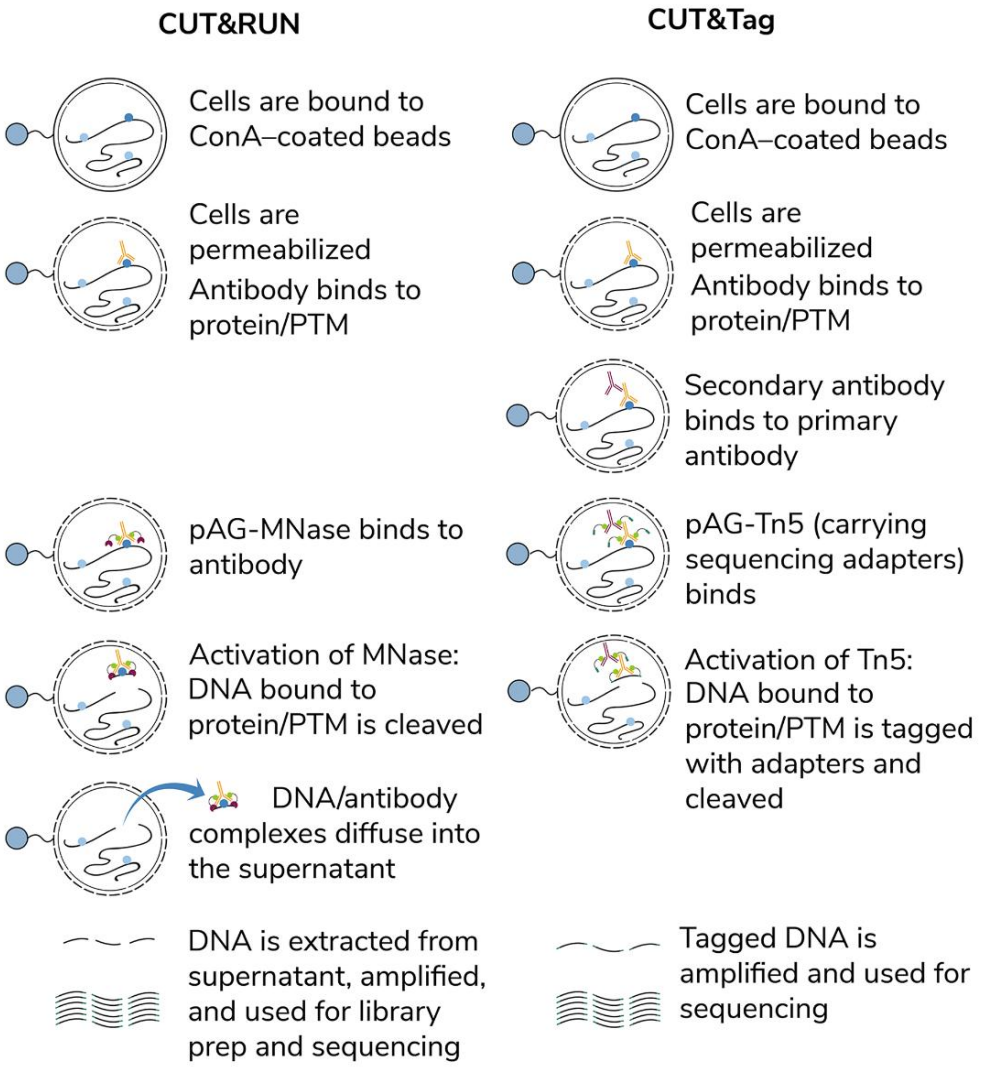
- Cleavage Under Targets and Tagmentation
- **pAG-Tn5**
- Eliminate the need for chromatin fragmentation and library preparation.

Pros:

- more time efficient

- amenable to **lower sample inputs and seq depth**

- particularly notable in single cell inputs



Cut&RUN or Cut&TAG:

CUT&RUN outperforms CUT&Tag:

- **Profiling transcription factor binding**

- **Map at higher resolution** due to the smaller size of the MNase protein and nature of its reaction

CUT&Tag offers a benefit over CUT&RUN:

- **Elimination of library preparation steps (?)**

- **Reduces time, resources, and cell input requirements**

Comparisons

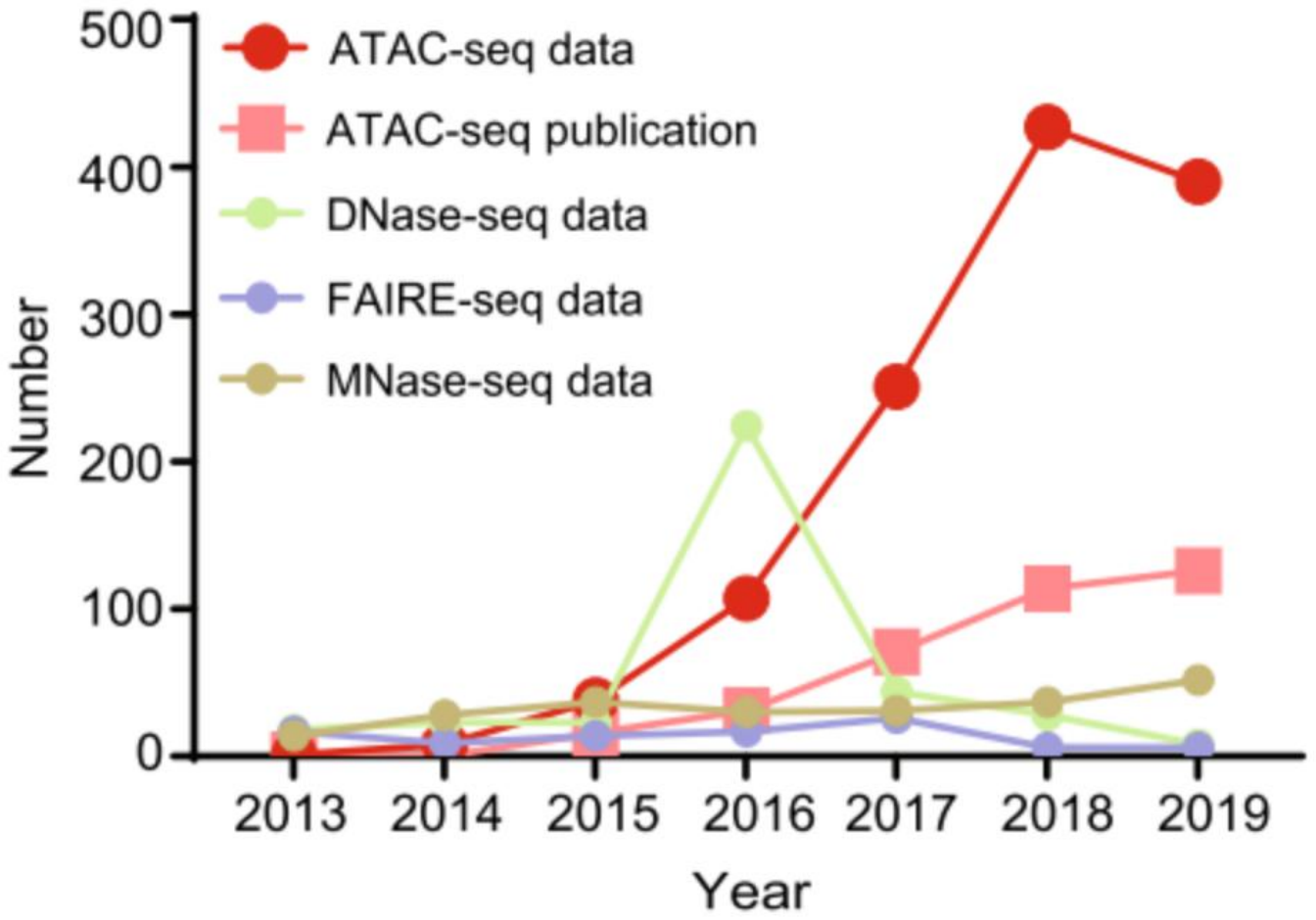
Features	ChIP-Seq	CUTANA™ CUT&RUN	CUTANA™ CUT&Tag
Sample Input	Sheared Chromatin	Cells OR nuclei	Nuclei (recommended)
Typical Required Cell #	> 1 Million	500K	100K
Ideal Targets	Histone PTMs & chromatin-interacting proteins	Histone PTMs & chromatin-interacting proteins, including remodelers	Histone PTMs & select validated targets
Secondary Antibody	No	No	Yes
Library Preparation	Yes	Yes	No (Direct to PCR)
Protocol Time (Cells → NGS libraries)	~ 1 week	2 days (can be automated)	2 days (can be automated)
Sequencing Depth	> 30 million	3-5 million	3-5 million
Signal : Noise	Low	High	High
Experimental Throughput	Low	High	High

Table 1: Comparison of ChIP-seq, CUT&RUN, and CUT&Tag.

DNA accessibility identification

Part3

Publication number



3.1 DNA accessibility: DNase–seq

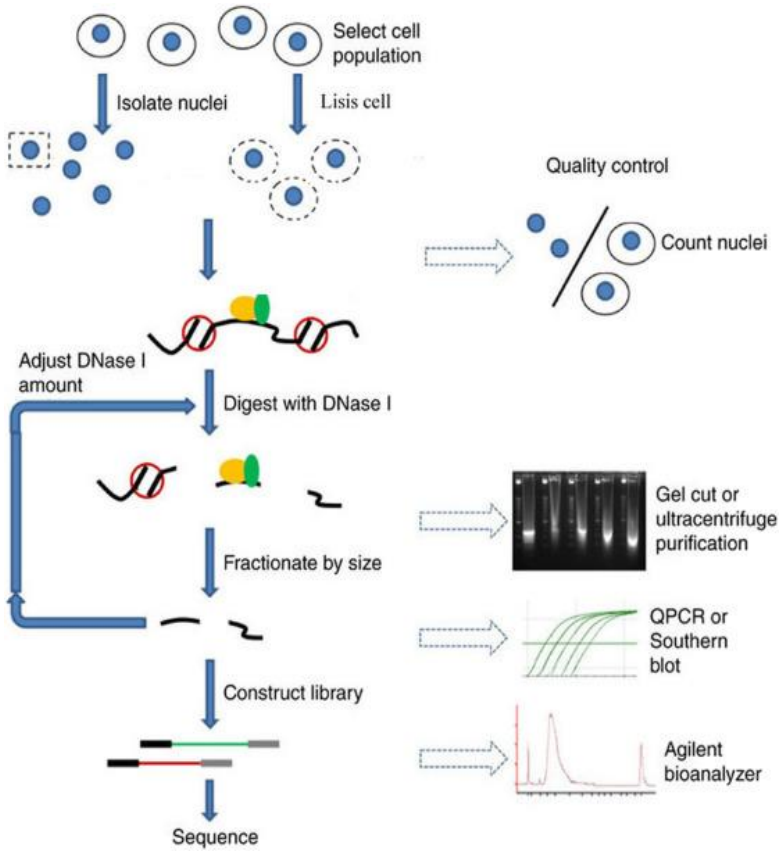
- Digestion by the **DNaseI nuclease** to identify regions of nucleosome-depleted **open chromatin** where there are **binding sites** for all types of factors

Pros:

- Comprehensive peaks
- Higher peak resolution (often < 200bp) compared with FAIRE-seq

Cons:

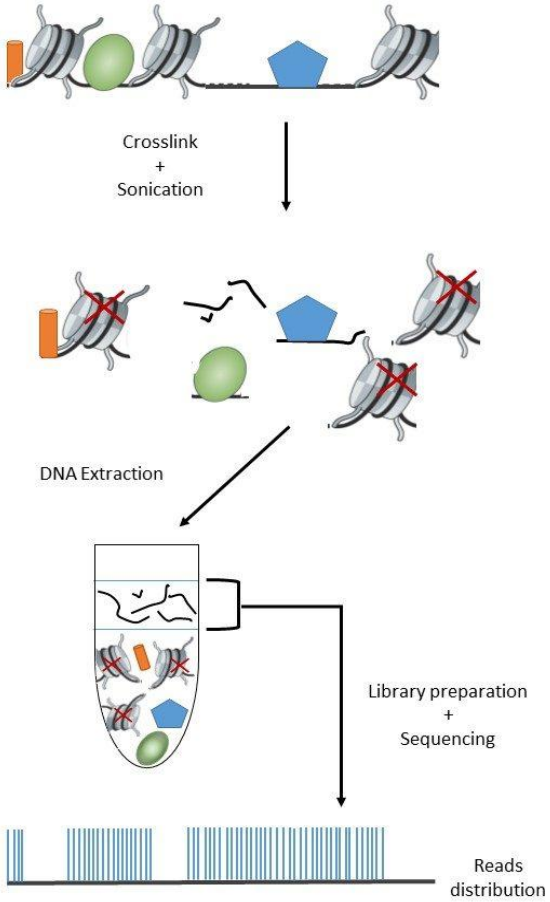
- DNA loss through the multiple purification steps limits sensitivity.



Song and Crawford, 2010

3.2 DNA accessibility: FAIRE–seq

- Formaldehyde-assisted identification (FAIRE–seq)
- Identifies nucleosome-depleted regions by extracting fragmented DNA that is not crosslinked to nucleosomes.



Pros:

Simple and highly reproducible protocol

- Does not require antibodies, enzymes (DNase or MNase), single-cell suspension or nuclear isolation

- Easily adapted for use on tissue samples

Cons:

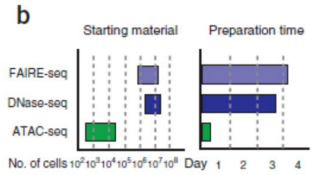
- Background noises

- Low resolution rate

- Time control of Formaldehyde-assisted

3.3 DNA accessibility: ATAC-Seq

- Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-Seq)
- Use the hyperactive transposase **Tn5**



Pros:

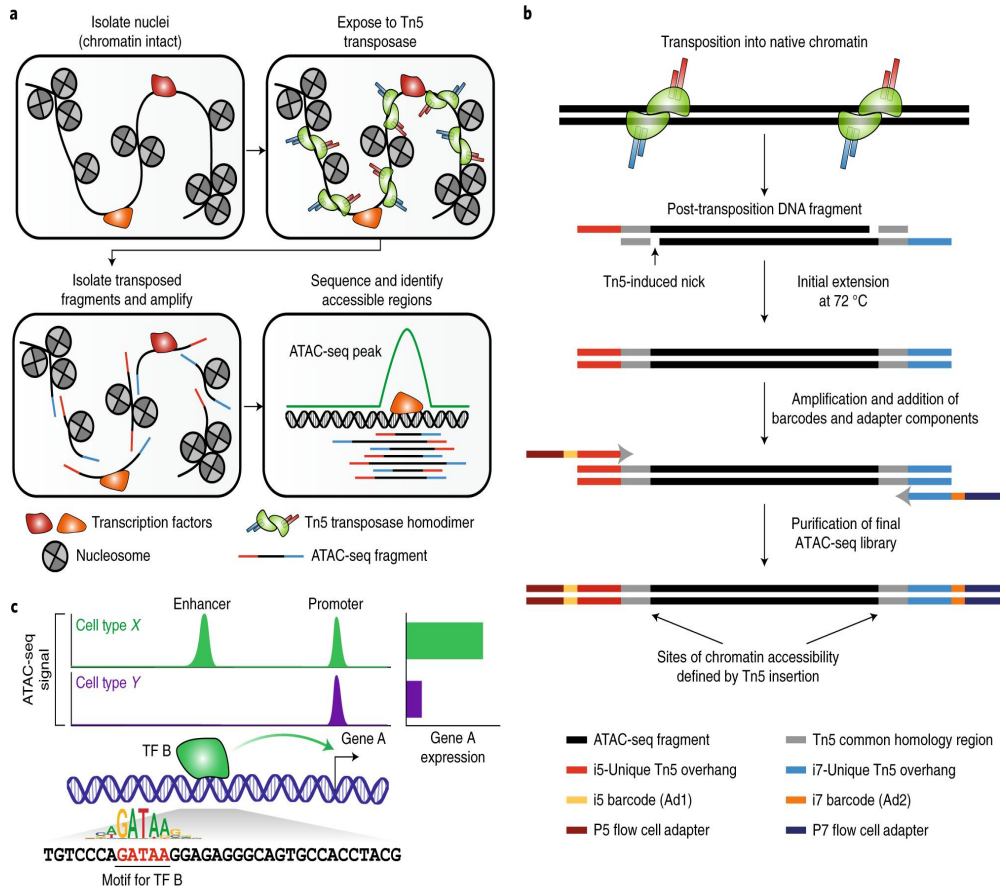
- **Easy and fast experimental techniques**
- Require **less starting material (cell num)**

(e.g. tissues with small cell count)

Cons:

- potential contamination of mitochondrial DNA
- potential sequence or structural biases

of Tn5 enzyme



Buenrostro et al. 2013

3.3 DNA accessibility: ATAC-Seq

Critical Parameters:

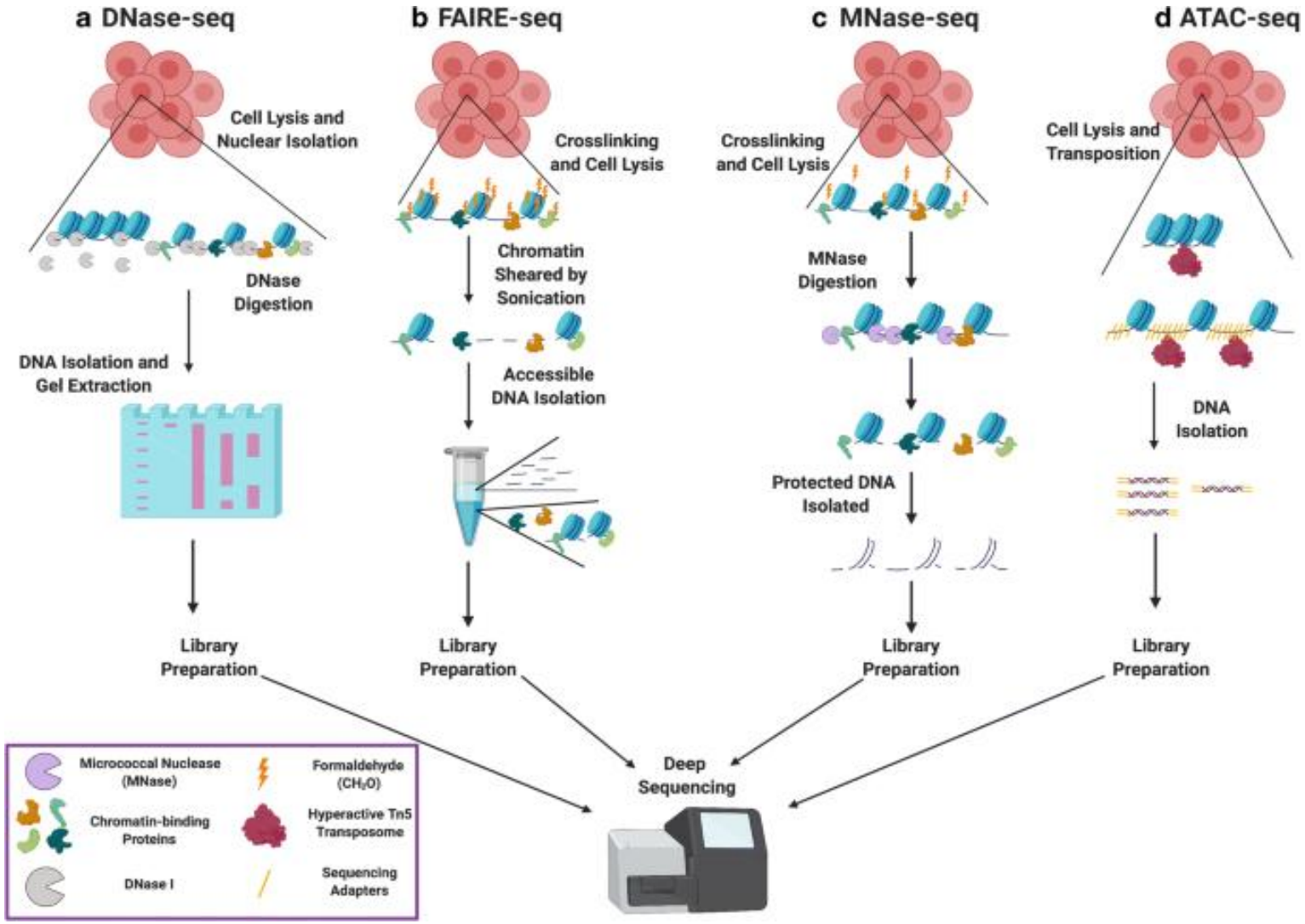
- 1. **Cell number:** diversity of the library (unique fragments) & noise & difficult for sequencing

- 2. **Fragment size/distribution:** a window of 100–1000 bps(large size) to maintain high library complexity, and enable the richness of the inferences that can be extracted from the full fragment size distribution.

Required cell num and variance of fragment

- - Depending on the **species**
- - If results not ideal, scale the reaction down 10x and to optimize the lysis methods and the cell num, starting from 5,000 cells per reaction.

Summary of DNA accessibility

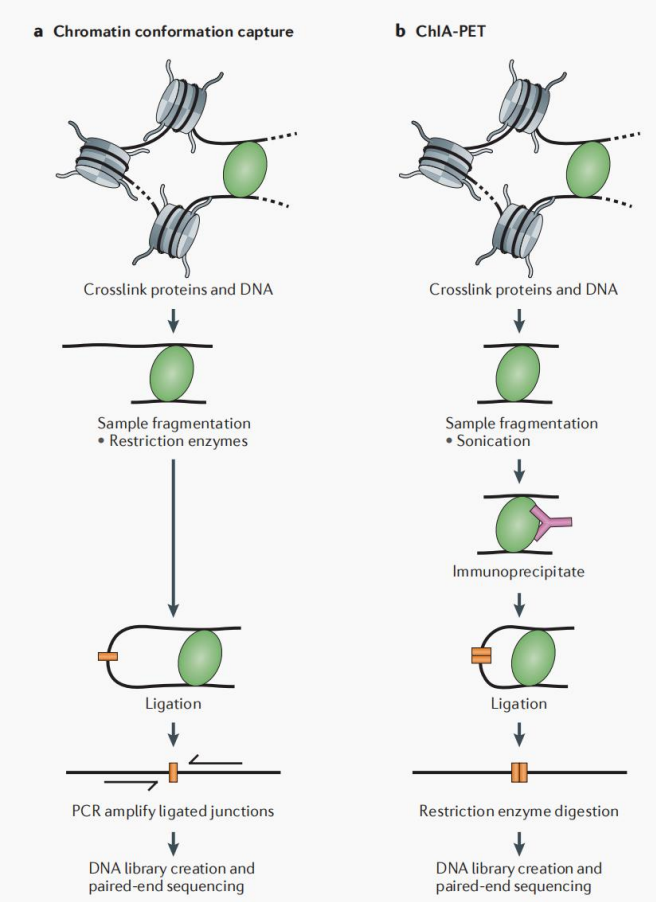


Klein, Chromosome Research, 2020

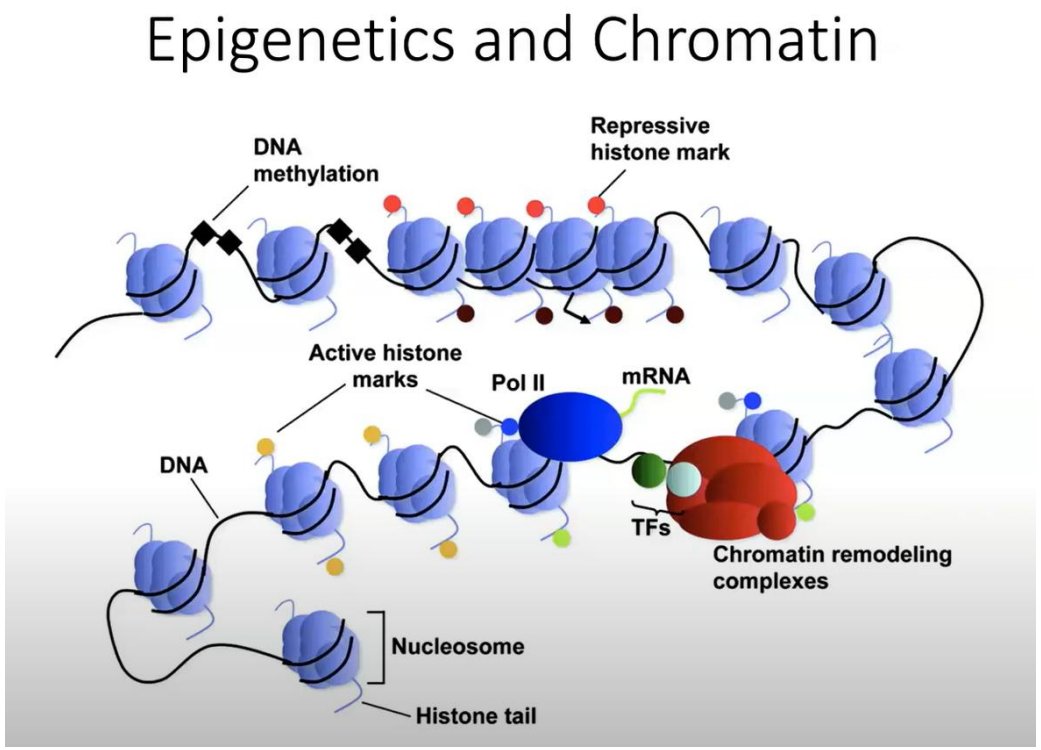
3.4 Chromatin interaction

Chromatin conformation capture & ChIA-PET

Regulate gene expression by bringing distal regulatory elements, such as **super-enhancers**, to promoters in close spatial proximity



Others



- Epigenetics beacons in the genome
- Histone marks
  - Annotate functional elements in the genome
  - TF binding from nucleosome and chromatin dynamics
  - Find new lncRNA genes
  - Bivalent genes, super enhancer and super promoter targets
- Chromatin accessibility
  - DNase-seq and ATAC-seq: collection of TF binding sites
  - Identify driving TFs and active genes in a biological process
  - Caution of modeling “footprint”