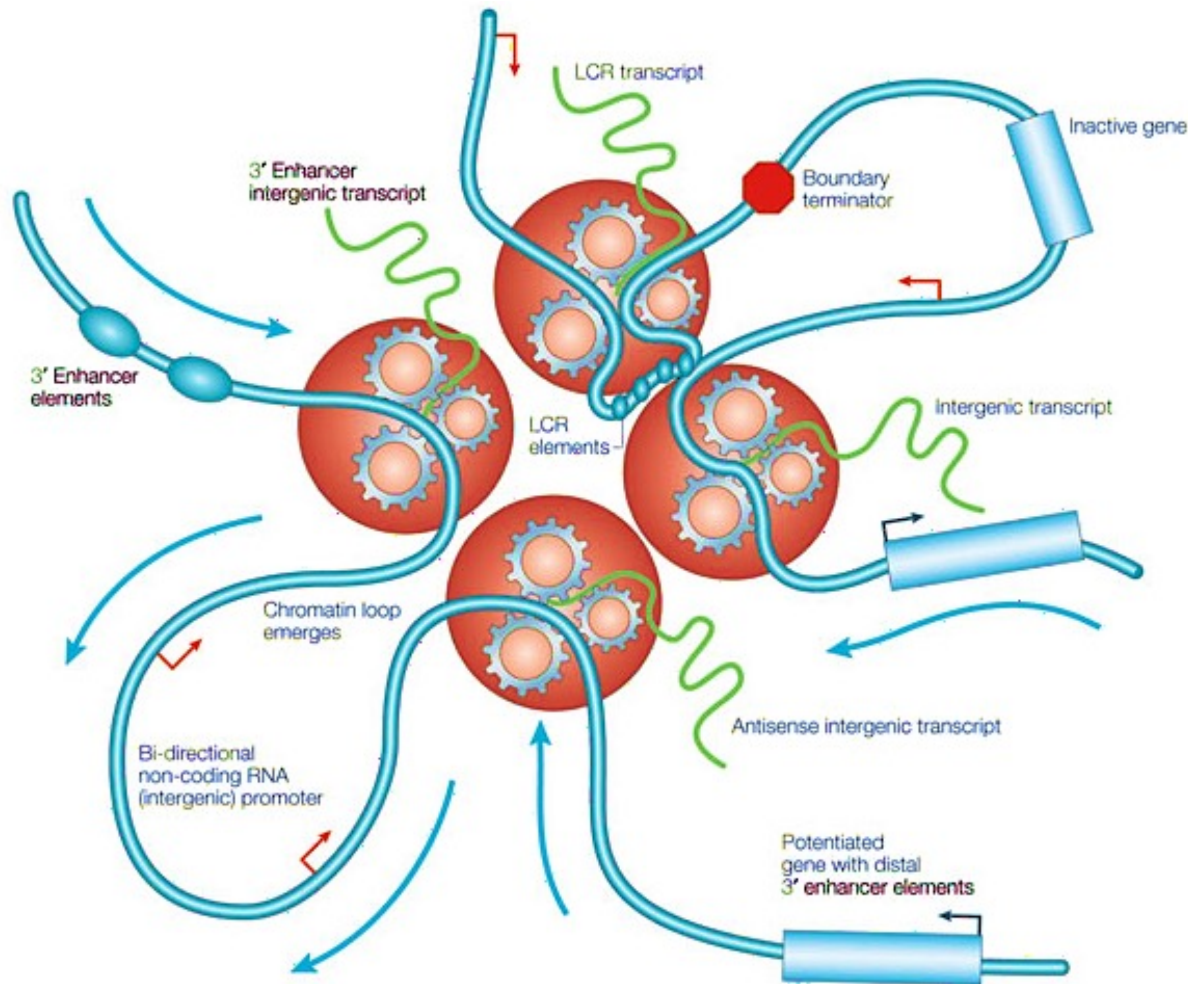
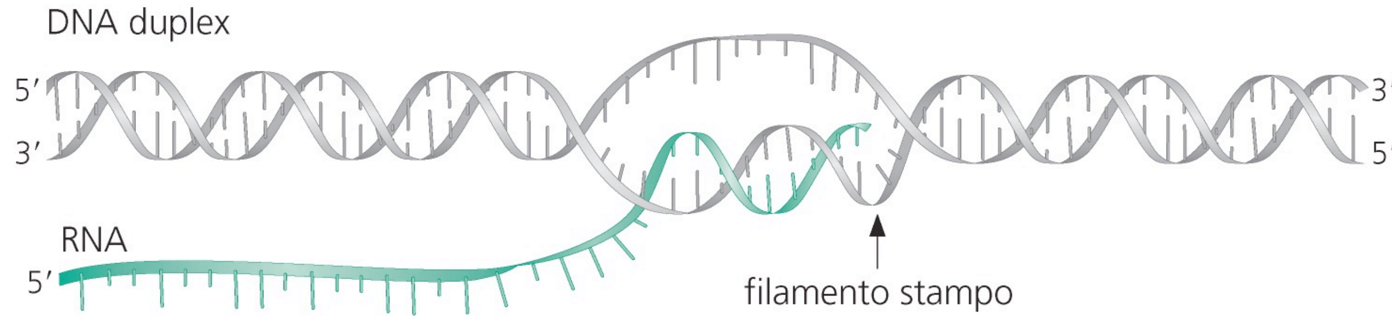
GENE EXPRESSION REGULATION IN EUKARYOTES



Chakalova *et al*. *Nature Rev Genetics* 6, 669-677.

Chromatin structure and *Transcriptional* regulation_I part

Transcription is the *first* step of gene expression and it is the process by which the information contained into DNA is *converted* in RNA.



Eukaryotic Transcriptional Regulation

1. Level of Chromatin (*DNA accessibility*)

- Histone modifications
- Nucleosome composition
- DNA methylation

2. Level of DNA (*Interaction with basal transcription machinery*)

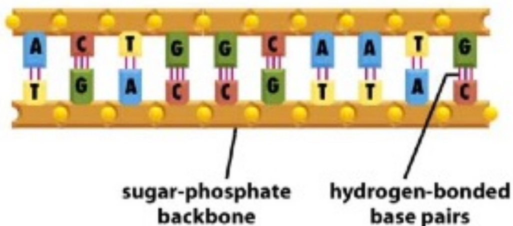
- Regulatory sequences (enhancers, silencers)
- Transcription factors (activators, repressors)

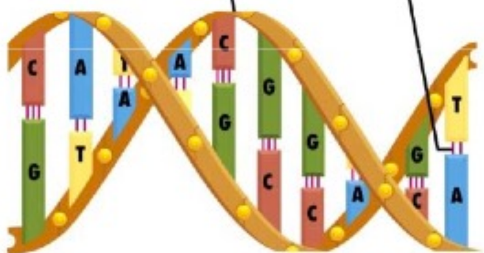
3. Level of *Regulatory RNA* (Interaction with DNA, RNA or protein)

- *small* and *long* noncoding RNAs

The problem: DNA lenght is always higher than the dimension of the compartment in which it is stored

Compartment	Shape	Dimensions	Type of nucleic acid	Length
TMV	filament	0.008 × 0.3 μm	1 single-stranded RNA	2 μm = 6.4 kb
Phage ϕd	filament	0.006 × 0.85 μm	1 single-stranded DNA	2 μm = 6.0 kb
Adenovirus	icosahedron	0.07 μm diameter	1 double-stranded DNA	11 μm = 35.0 kb
Phage T4	icosahedron	0.065 × 0.10 μm	1 double-stranded DNA	55 μm = 170.0 kb
E. coli	cylinder	1.7 × 0.65 μm	1 double-stranded DNA	1.3 μm = 4.2×10^3 kb
Mitochondrion (human)	oblate spheroid	3.0 × 0.5 μm	~10 identical double-stranded DNAs	50 μm = 16.0 kb
Nucleus (human)	spheroid	6 μm diameter	46 chromosomes of double-stranded DNA	1.8 m = 6×10^6 kb



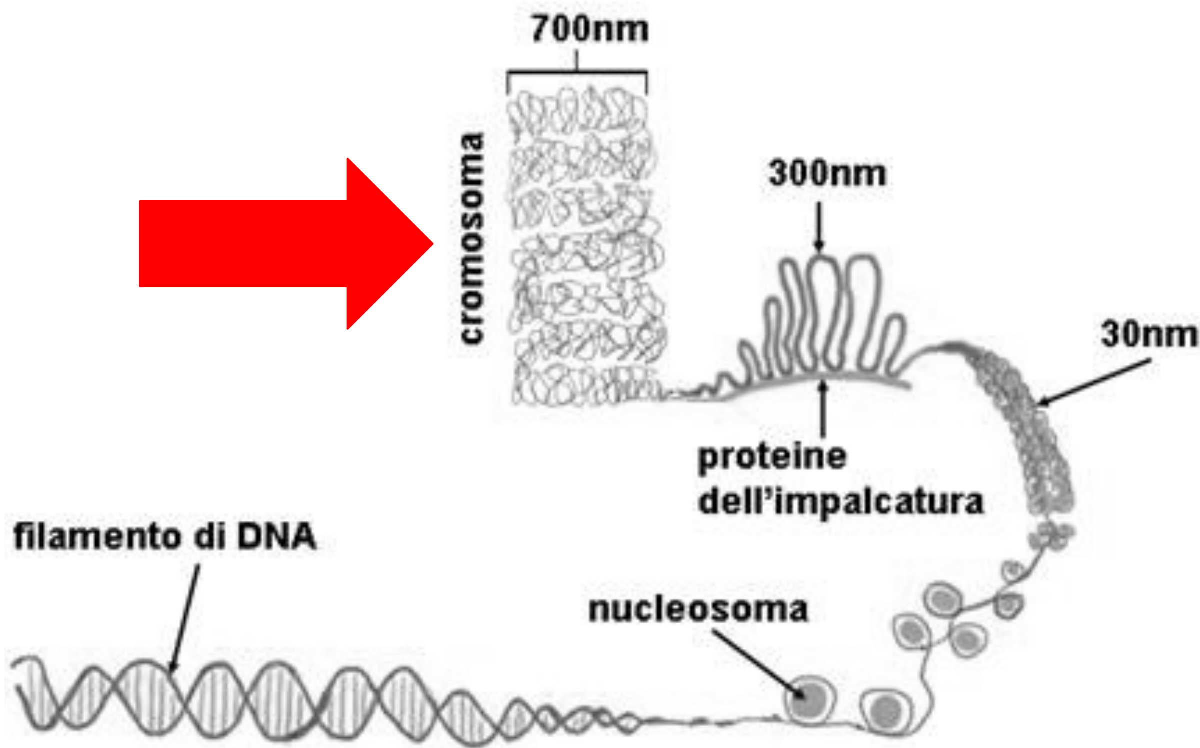


- Spacing between base pairs ≈3.4Å
- For human genome, approximately 3.2 billion base pairs
- Total length ≈ $3.4 \times 10^{-10} \times 3.2 \times 10^{9} \times 2 \approx 2.2$m
- Diameter of a nucleus: $5 \sim 10 \times 10^{-6}$m

Packaging is essential for several reasons:

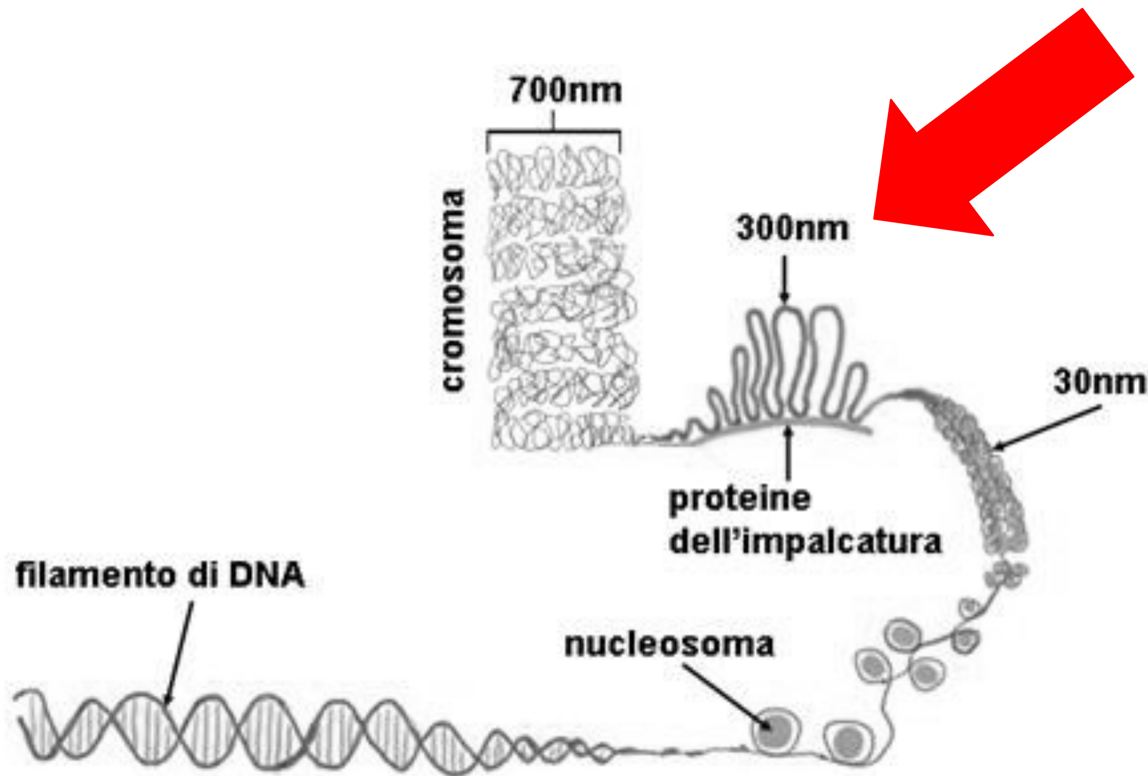
- DNA molecules would be ***too long*** to fit inside cells
- ***Damage*** protection
- during cell division, it is essential that DNA remains intact and evenly distributed among cells. Chromosomes are a key part of the process that ensures DNA is accurately *copied* and distributed in the vast majority of cell divisions.

Let’s sum-up



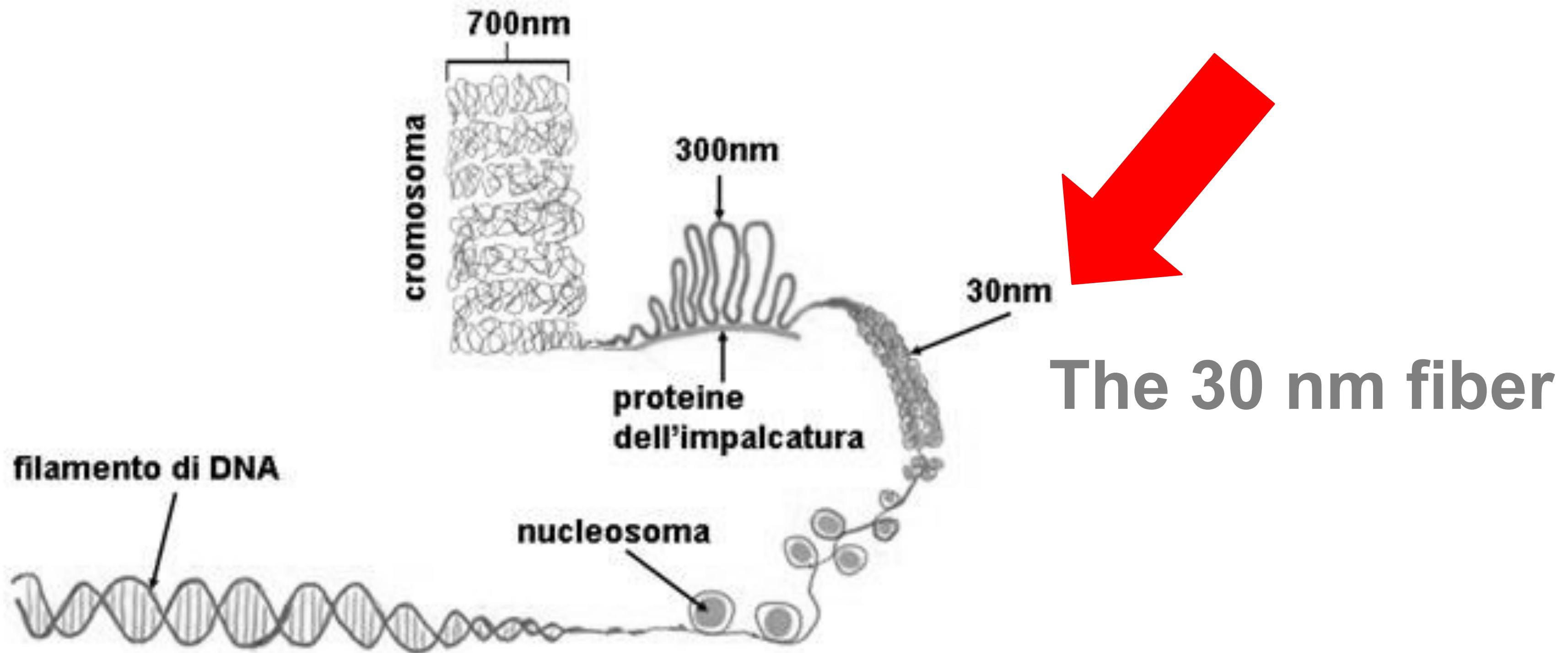
Chromatin is a nucleoprotein complex whose primary function is to pack DNA and to organize eukaryotic genomes.

Let's sum-up



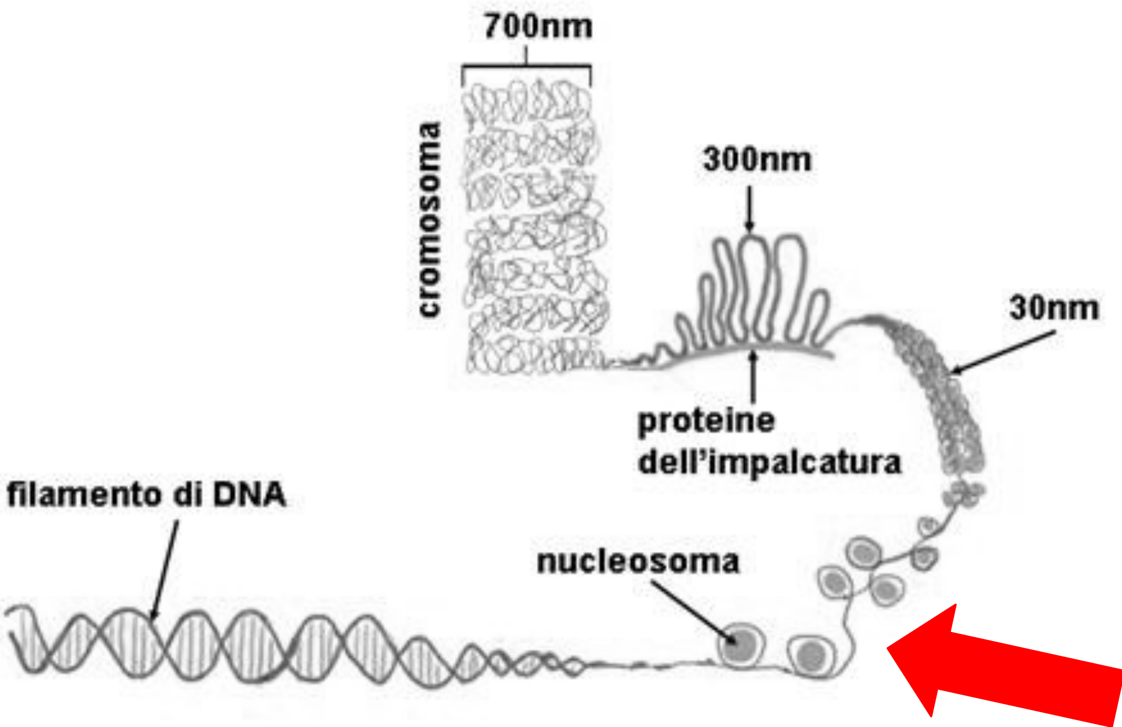
Higher-order structures: the chromatin loops

Let’s sum-up

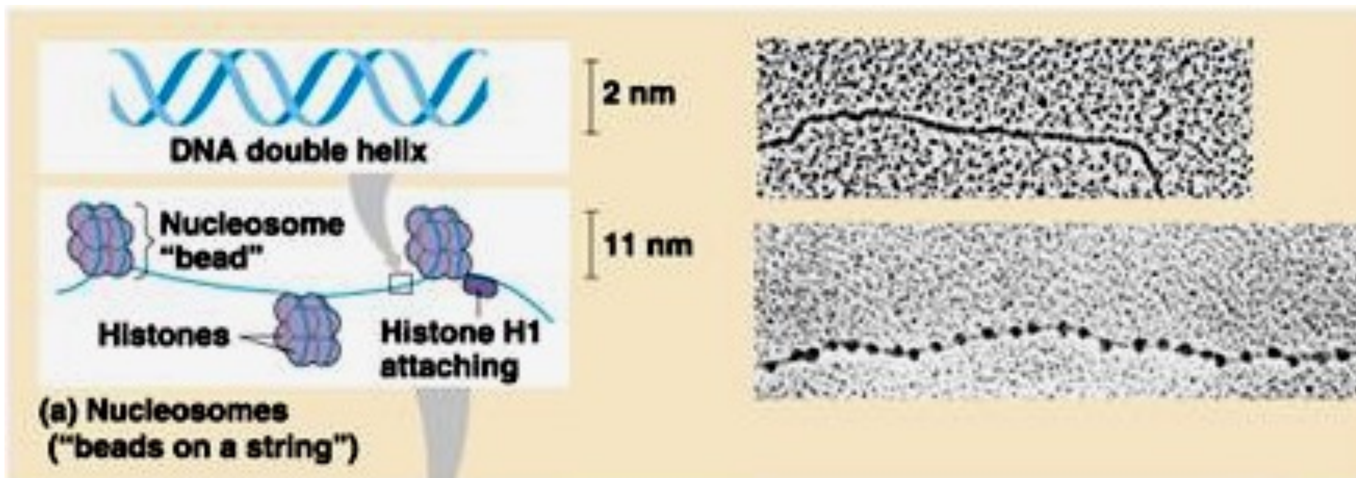


The 30 nm fiber

Let's sum-up



The 10 nm fiber

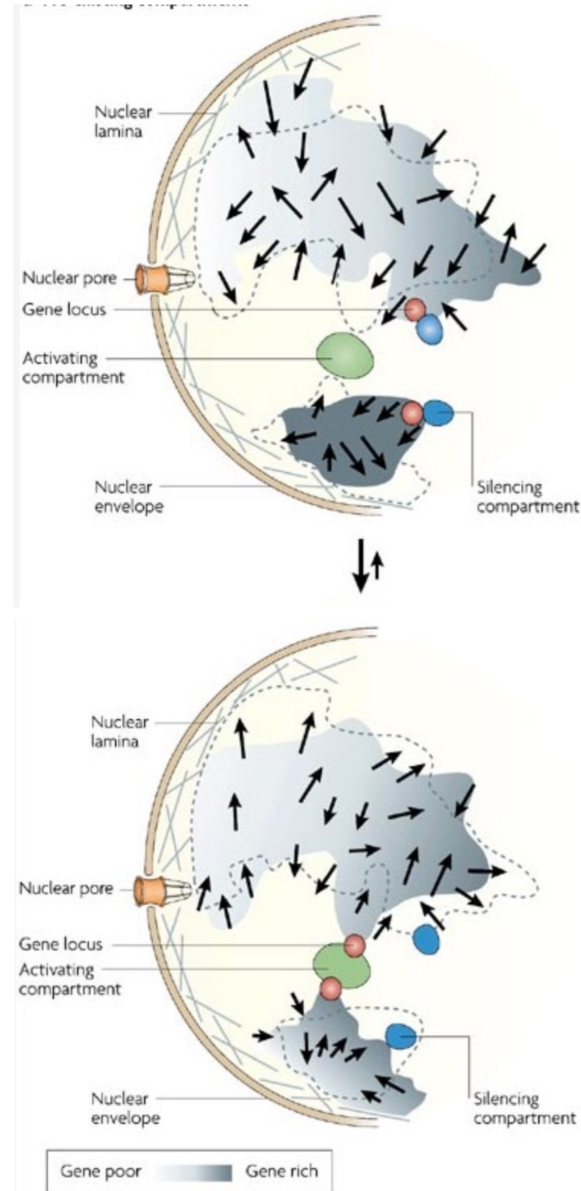


The observation by electron microscopists that chromatin appeared similar to **beads** on a string

Chromatin is an highly dynamic structure

Chromatin can no longer be considered as merely the sum of independent regions but rather should now be considered as a *flexible* and *interconnected* web in which neighbouring, as well as distant, domains can interact

Chromatin mobility allows **dynamic interactions** between genomic loci and between loci



Gene **repression**

Gene **activation**

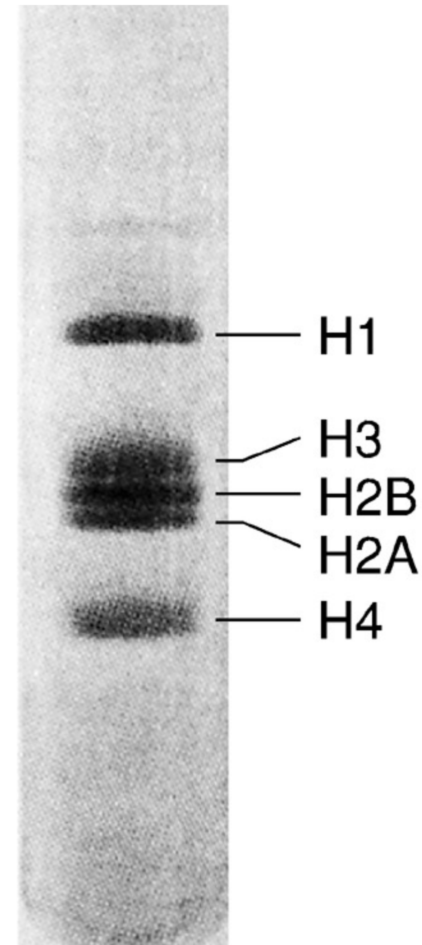
FROM THE FOLLOWING ARTICLE:
Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions
Christian Lanctôt, Thierry Cheutin, Marion Cremer, Giacomo Cavalli & Thomas Cremer
Nature Reviews Genetics **8**, 104-115 (February 2007)

Let's sum-up

Packaging is accomplished via *highly conserved proteins* called **histones**, which are central components of chromatin.

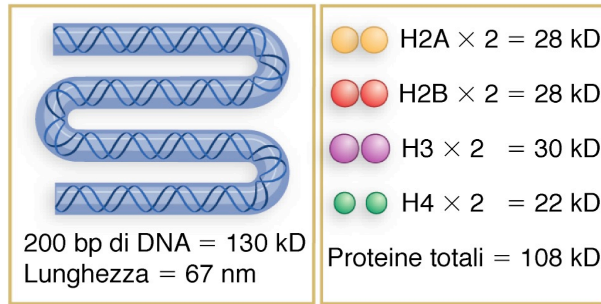
Histones are a family of *small, positively* charged proteins (Van Holde, 1988).

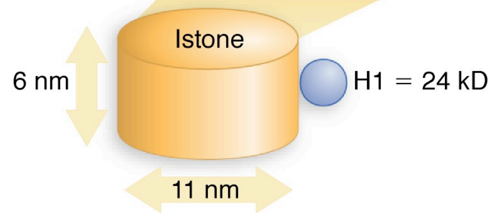
As DNA is negatively charged, due to the phosphate groups in its phosphate-sugar backbone, histones bind with DNA very tightly.

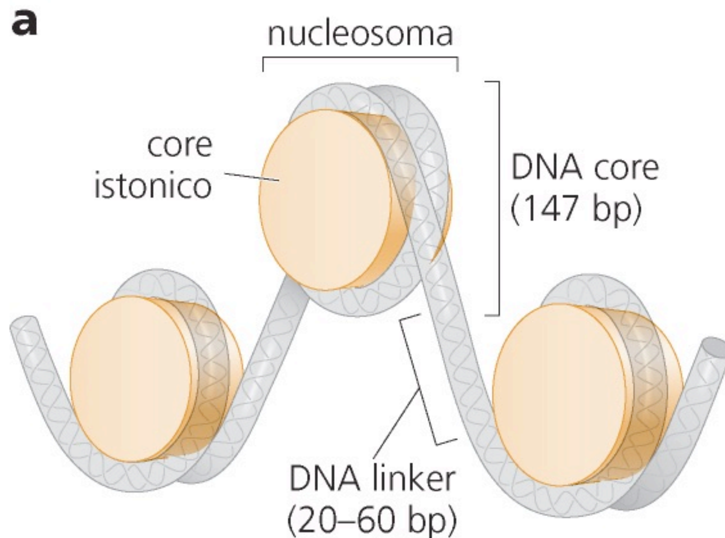


chem. & Biophys. 130, 1969, f. 6A, p. 343. Reprinted with

Nucleosome is the fundamental *repeating* unit of chromatin





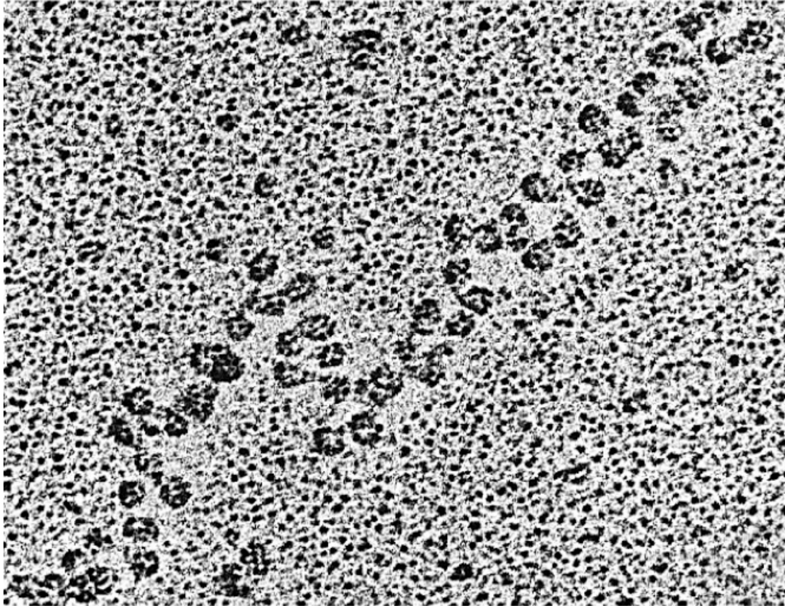


- two each of the histones **H2A**, **H2B**, **H3**, and **H4** come together to form a **histone octamer**, which binds and wraps approximately 1.7 turns of DNA, or about **146** base pairs.

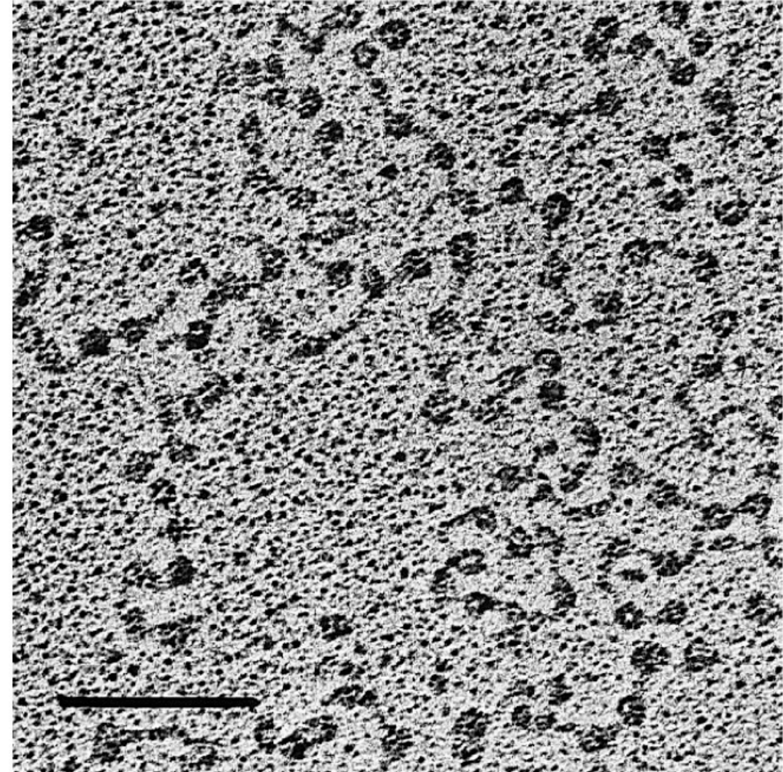
- the *addition* of one **H1** protein *wraps another 20 base pairs*, resulting in two full turns around the octamer, and forming a structure called a *chromatosome.* This joining DNA is referred to as *linker* DNA.

Effect of H1 absence on nucleosomes packaging (Electron Microscopy)

With H1

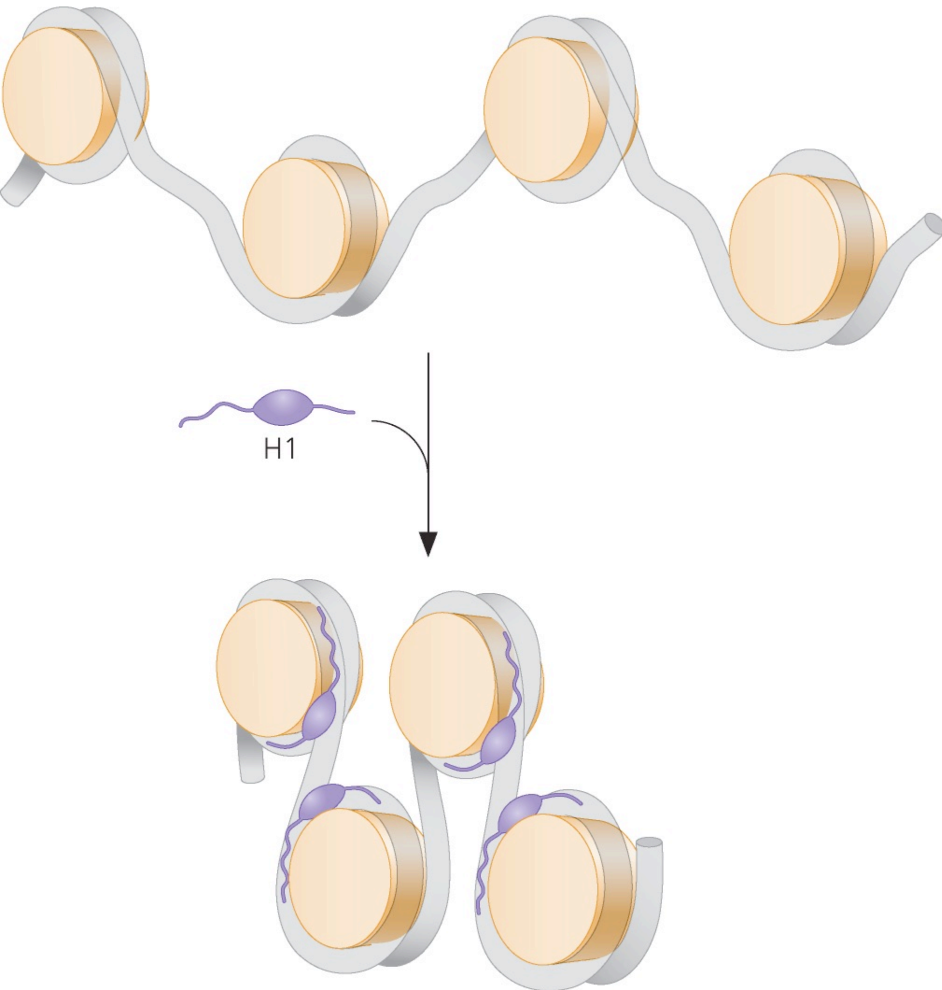


Without di H1



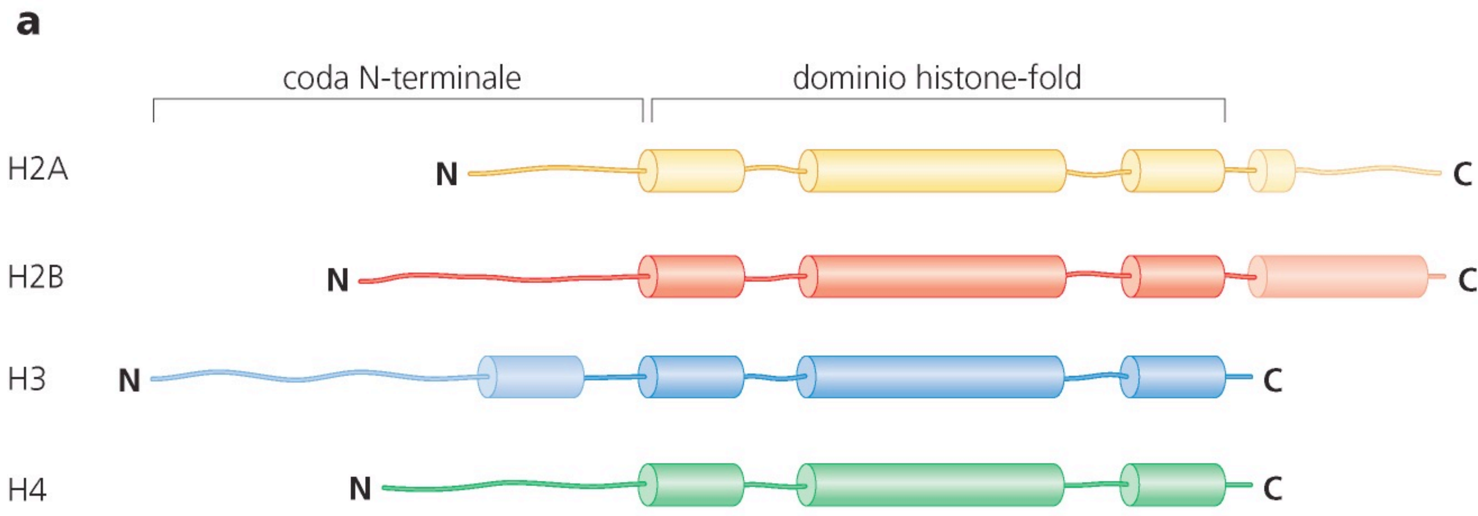
Watson et al., BIOLOGIA MOLECOLARE DEL GENE, Zanichelli editore S.p.A. Copyright © 2005

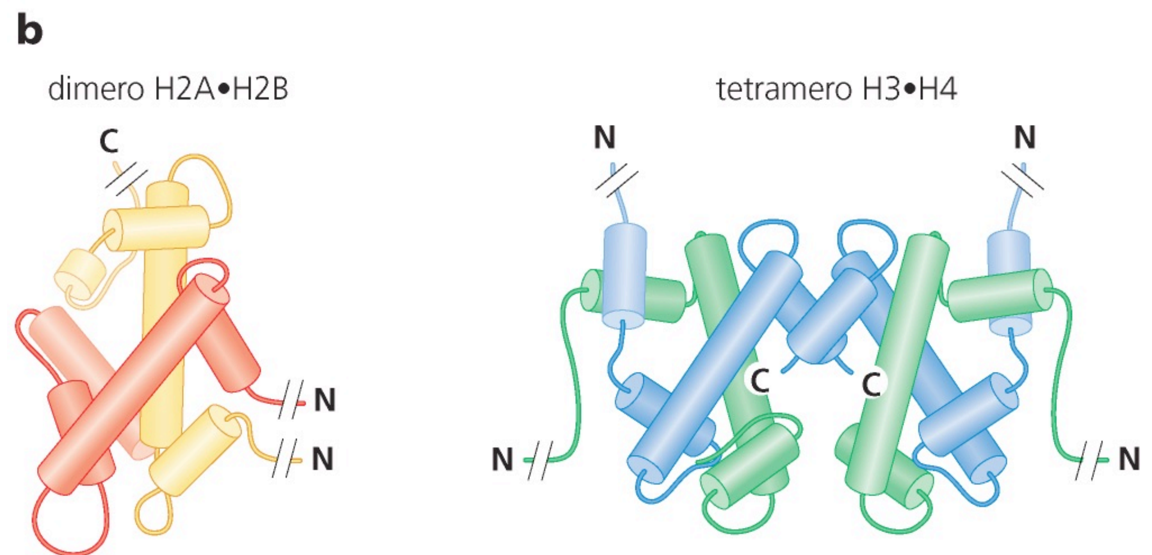
With the help of the linker **histone H1** that binds the linker DNA connecting nucleosomes, chromatin forms *higher-order structures* that enable eukaryotic cells to accommodate and organize genomic DNA inside their nucleus.



- **H1** is outside the nucleosome and binds DNA at the level of entry/exit points
- **H1** contributes to the formation of higher ordered structures: the 30 nm fiber.

Histones consist of a flexible tail and a globular core domain that folds into the characteristic histone fold. The ***N-term tails*** protrude outside and interact with DNA





Different ways to *modify* the chromatin exists which regulates *gene expression*, DNA repair, replication and recombination.

1)Histone variants (i.e. H2A e H3)

2)Post-translational histone modifications (PTM): acetilation, methylation etc….

3)DNA modifications

4)Non coding RNA

1) Histone variants

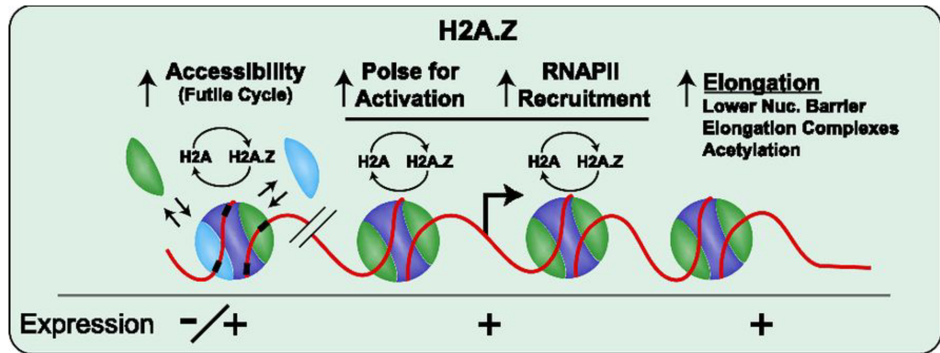
- Histone variants have ***distinct* amino acid sequences** that can influence both the physical properties of the nucleosome and nucleosome dynamics.

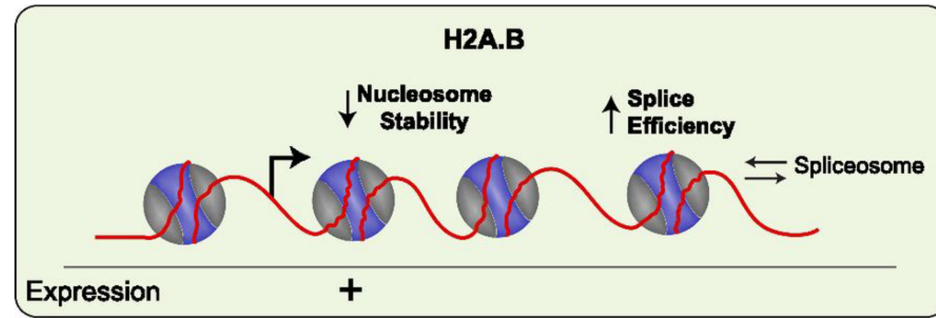
- Canonical histones are deposited in a *replication-coupled* manner to package the newly replicated genome. In contrast, histone variants are expressed *throughout the cell cycle* and *replace* canonical histones when nucleosomes are evicted

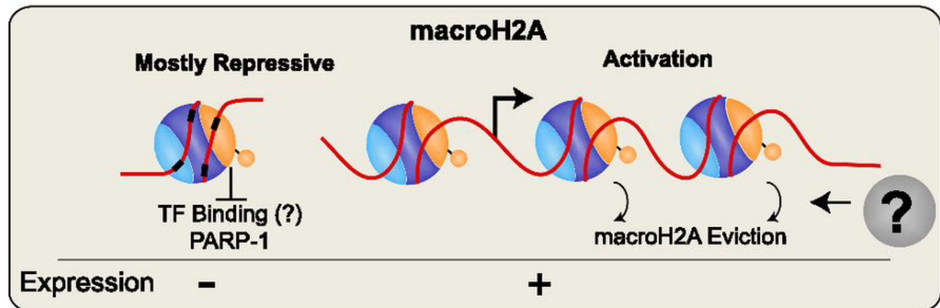
- histone variants alter nucleosome structure, stability, dynamics, and, ultimately, DNA accessibility.

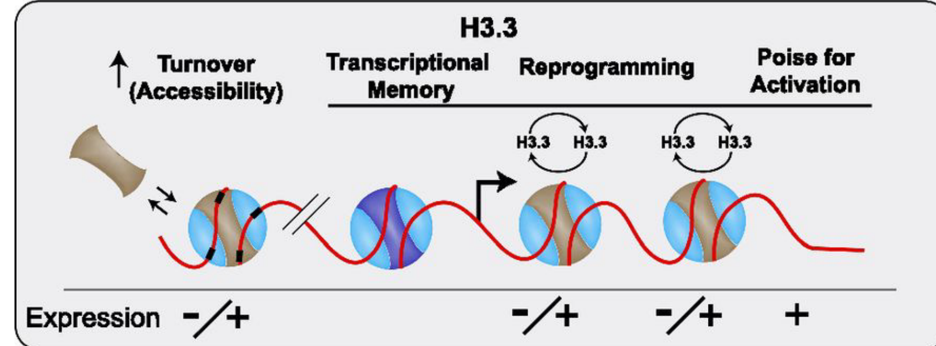
During transcription, histone variants shape the chromatin landscape of *cis*-regulatory and coding regions in support of specific transcription programs.

General role of histone variants and their deposition pathways on transcriptional regulation









Christopher M. Weber, and Steven Henikoff Genes Dev. 2014;28:672-682

H2A variants are the most diverse, perhaps reflective of relaxed structural constraint within the nucleosome.

One such variant, H2A.Z, is only ~60% identical to H2A

Different ways to *modify* the chromatin exists which regulates gene expression, DNA repair, replication and recombination.

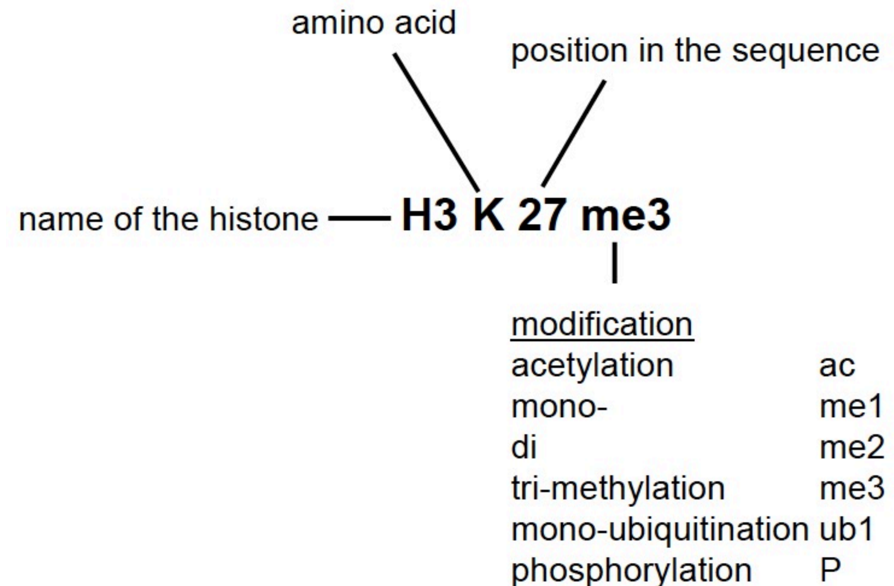
1)Histone variants (i.e. H2A e H3)

2)Post-translational histone modifications (PTM): acetilation, methylation etc.…

3)DNA modifications

4)Non coding RNA

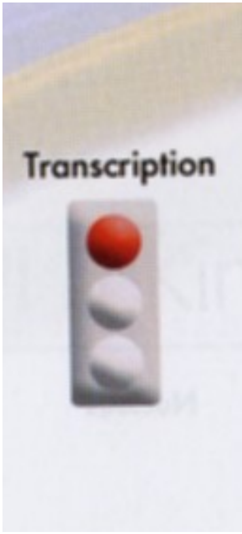
Nomenclature

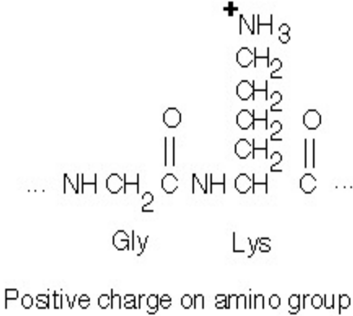


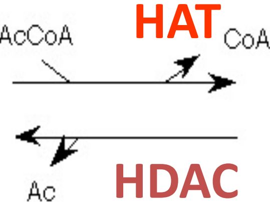
2) Post-translational histone modifications (PTM)

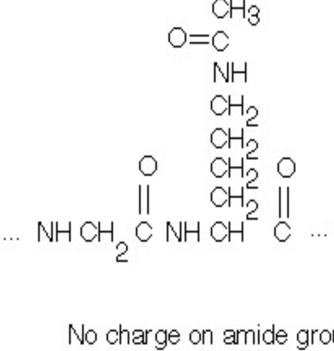
PTM have a number of different functions

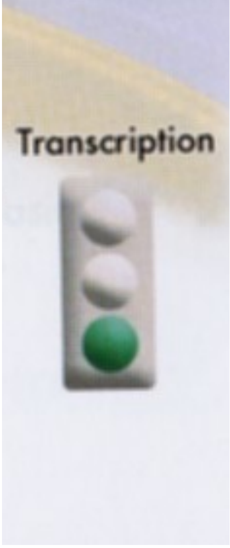
- histone modifications result in a *change* in the net *charge* of nucleosomes, which could loosen inter- or intranucleosomal DNA-histone interactions. For instance, *acetylated histones are easier to displace from DNA!!!*









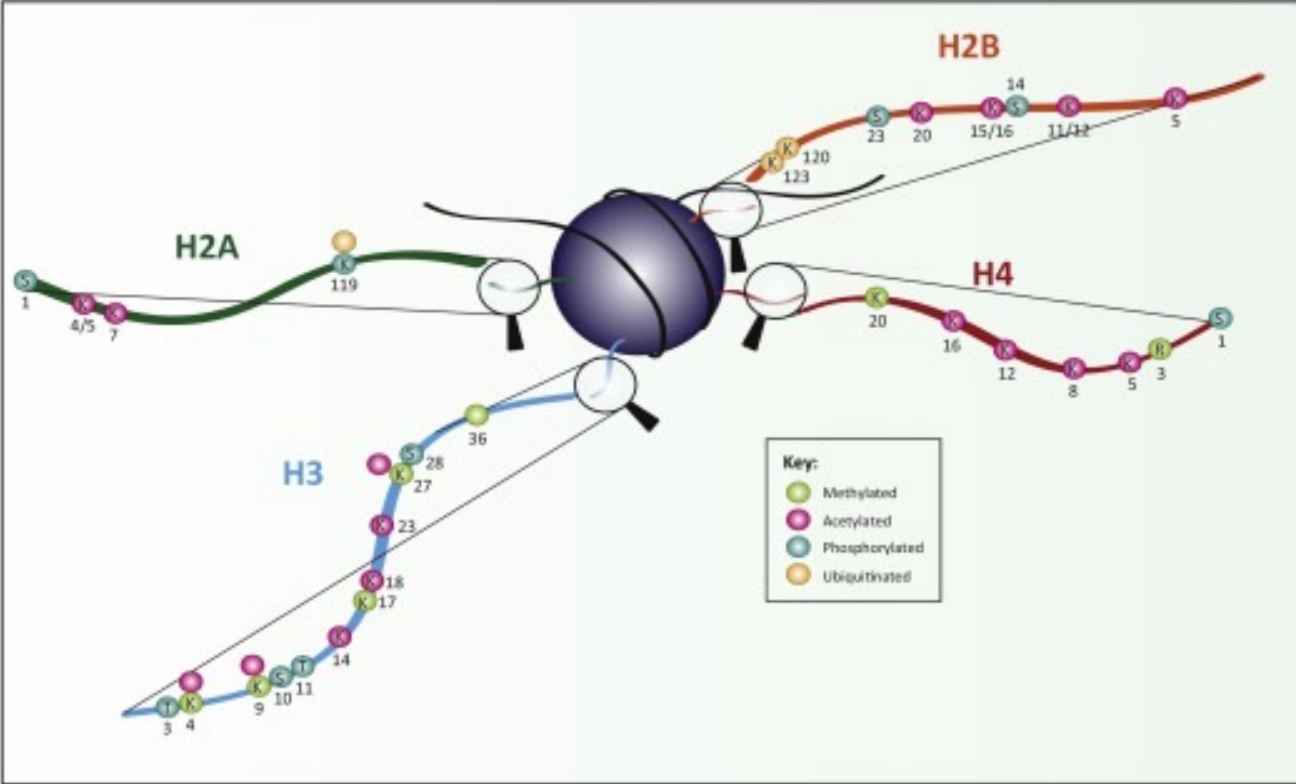


- individual histone modifications or modification patterns are read by other proteins that influence chromatin dynamics and function

- some modifications directly influence higher-order chromatin structure.
  H4 K16 inhibits the formation of compact 30 nm fibers

2) Post-translational histone modifications (PTM)

The best-studied modifications are those occurring on the N-terminal 'tail' regions of the histones, which project from the nucleosome and are accessible on its surface



Trends in Genetics

These modifications include:

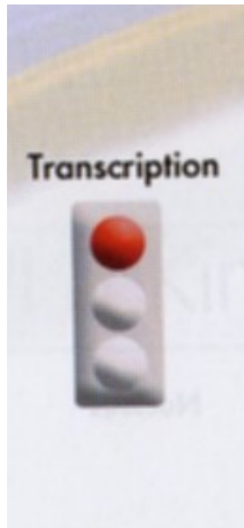
- Acetylation
- Methylation
- Phosphorylation
- Ubiquitinylation
- Sumoylation
- ADP ribosylation
- Deamination

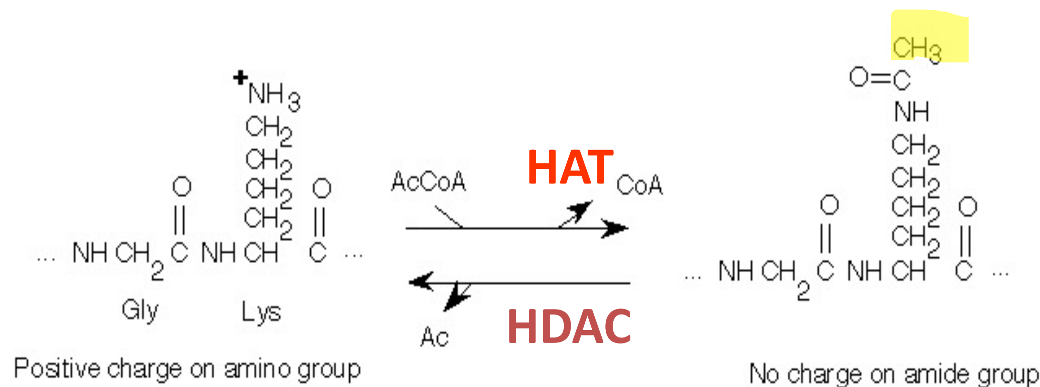
Lateral Thinking: How Histone Modifications Regulate Gene Expression

Moyra Lawrence, Sylvain Daujat, Robert Schneider

PTM have several functions

- histone modifications result in a *change* in the net *charge* of nucleosomes, which could loosen inter- or intranucleosomal DNA-histone interactions. For instance, ***acetylated*** *histones are easier to displace from DNA!!!*





- individual histone modifications or modification patterns are read by other proteins that influence chromatin dynamics and function

- some modifications directly influence higher-order chromatin structure.
  H4 K16 inhibits the formation of compact 30 nm fibers

Covalent modifications of N-term tails: NOMENCLATURE

1) Histone acetylases (HATs)
2) Histone de-acetylases (HDACs)
3) Histone methylases (HMTs)
4) Histone de-methylases (HDMs)
5) Histone ubiquitin ligases

Some of the modifications in these tails can directly affect the *interactions* between nucleosomes.

Histone tail modifications can increase DNA compaction; for example, **H4K20** di- and tri-methylation, which have been shown to *enhance* chromatin condensation

The **H4K16ac** has been shown to *reduce* chromatin compaction and increase transcription

Histone	Modification	Role
H4	H4R3me	Transcriptional activation
	H4K20me1	Transcriptional silencing
	H4K20me3	Heterochromatin
	H4K5ac	Histone deposition; transcriptional activation; DNA repair
	H4K8ac	Transcriptional activation; DNA repair; transcriptional elongation
	H4K12ac	Histone deposition; telomeric silencing; transcriptional activation; DNA repair
	H4K16ac	Transcriptional activation; DNA repair
	H4S1P	Mitosis

Table 1. Histone Tail Modifications

Histone	Modification	Role
H2A	H2AS1P	Mitosis; chromatin assembly
	H2AK4/5ac	Transcriptional activation
	H2AK7ac	Transcriptional activation
	H2AK119P	Spermatogenesis
	H2AK119uq	Transcriptional repression
H2B	H2BS14P	Apoptosis
	H2BS33P	Transcriptional activation
	H2BK5ac	Transcriptional activation
	H2BK11/12ac	Transcriptional activation
	H2BK15/16ac	Transcriptional activation
	H2BK20ac	Transcriptional activation
	H2BK120uq	Spermatogenesis/meiosis
	H2BK123uq	Transcriptional activation
H3	H3K4me2	Permissive euchromatin
	H3K4me3	Transcriptional elongation; active euchromatin
	H3K9me3	Transcriptional repression; imprinting; DNA methylation
	H3R17me	Transcriptional activation
	H3K27me3	Transcriptional silencing; X-inactivation; bivalent genes/gene poising
	H3K36me3	Transcriptional elongation
	H3K4ac	Transcriptional activation
	H3K9ac	Histone deposition; transcriptional activation
	H3K14ac	Transcriptional activation; DNA repair
	H1K18ac	Transcriptional activation; DNA repair; DNA replication
	H3K23ac	Transcriptional activation; DNA repair
	H3K27ac	Transcriptional activation
	H3T3P	Mitosis
	H3S10P	Mitosis; meiosis; transcriptional activation
	H3T11/S28P	Mitosis

Lateral Thinking: How Histone Modifications Regulate Gene Expression

Moyra Lawrence, Sylvain Daujat, Robert Schneider

Histone modifications associated with transcription

Modifications	Position		Enzymes *S. cerevisiae*	*S. pombe*	*Drosophila*	Mammals	Recognition Module(s)[a]	Functions in Transcription
Methylation	H3	K4	Set1	Set1	Trx, Ash1	MLL, ALL-1, Set9/7, ALR-1/2, ALR, Set1	PHD, Chromo, WD-40	Activation
		K9	n/a	Clr4	Su(var)3-9, Ash1	Suv39h, G9a, Eu-HMTase I, ESET, SETBD1	Chromo (HP1)	Repression, activation
		K27				E(Z)	Ezh2, G9a	Repression
		K36	Set2			HYPB, Smyd2, NSD1	Chromo(Eaf3), JMJD	Recruiting the Rpd3S to repress internal initiation
		K79	Dot1			Dot1L	Tudor	Activation
	H4	K20		Set9	PR-Set7, Ash1	PR-Set7, SET8	Tudor	Silencing
Arg Methylation	H3	R2				CARM1		Activation
		R17				CARM1		Activation
		R26				CARM1		Activation
	H4	R3				PRMT1	(p300)	Activation
Phosphorylation	H3	S10	Snf1				(Gcn5)	Activation
Ubiquitination	H2B	K120/123	Rad6, Bre1	Rad6		UbcH6, RNF20/40	(COMPASS)	Activation
	H2A	K119				hPRC1L		Repression
Acetylation	H3	K56					(Swi/Snf)	Activation
	H4	K16	Sas2, NuA4		dMOF	hMOF	Bromodomain	Activation
	Htz1	K14	NuA4, SAGA					Activation

[a] The proteins that are indicated within the parentheses are shown to recognize the corresponding modifications but specific domains have yet to be determined.

Polycomb and MLL/Trithorax Complexes

The ON and OFF states of key developmental genes are maintained by the **polycomb group (PcG)** and **MLL/Trithorax (Trx)** proteins, which mediate H3K27me3 to repress genes or H3K4me3 to activate genes.

Polycomb-group Proteins

- Maintains a silenced state
- Prevents chromatin remodelling

Trithorax-group Proteins

- Maintains an active state
- Counteracts the action of PcG proteins

Mol Cell. 2013 Mar 7;49(5):808-24. doi: 10.1016/j.molcel.2013.02.013.

Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put.

Simon JA[1], Kingston RE.

⊕ Author information

Abstract

Chromatin modification by Polycomb proteins provides an essential strategy for gene silencing in higher eukaryotes. Polycomb repressive complexes (PRCs) silence key developmental regulators and are centrally integrated in the transcriptional circuitry of stem cells. PRC2 trimethylates histone H3 on lysine 27 (H3K27me3), and PRC1-type complexes ubiquitylate histone H2A and compact polynucleosomes. How PRCs are deployed to select and silence genomic targets is the subject of intense investigation. We review advances on targeting, modulation, and functions of PRC1 and PRC2 and progress on defining the transcriptional steps they impact. Recent findings emphasize PRC1 targeting independent of H3K27me3, nonenzymatic PRC1-mediated compaction, and connections between PRCs and noncoding RNAs. Systematic analyses of Polycomb complexes and associated histone modifications during DNA replication and mitosis have also emerged. The stage is now set to reveal fundamental epigenetic mechanisms that determine how Polycomb target genes are silenced and how Polycomb silence is preserved through cell-cycle progression.

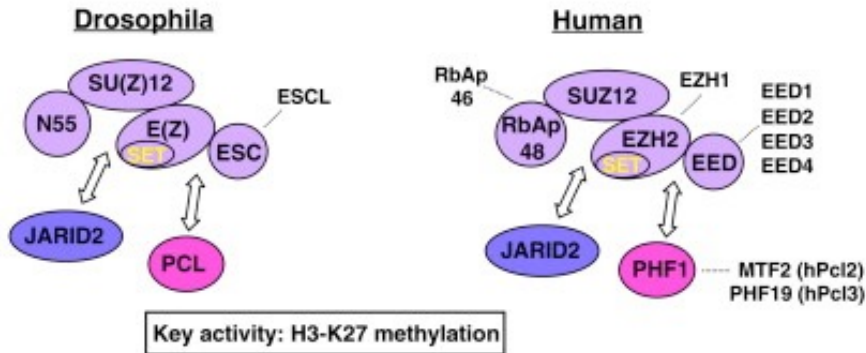
The central mechanistic question is:

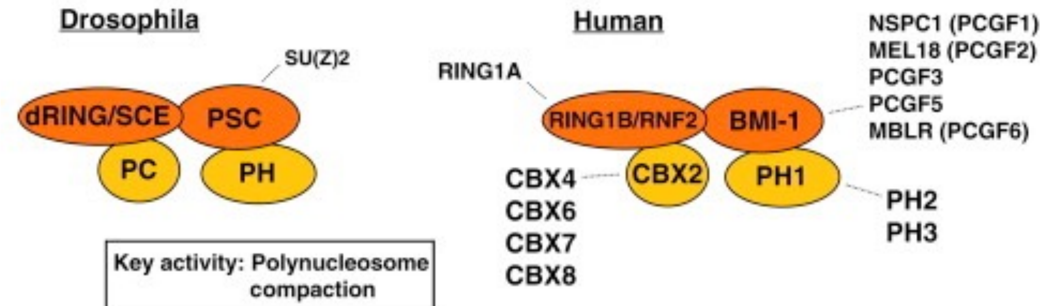
"How does PcG complexes shut down transcription?"

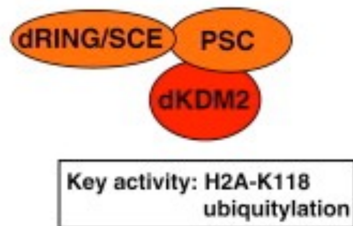
1) There is abundant literature on the histone modifications delivered by PcG complexes, **H3K27me3** and **H2AK119ub1**.

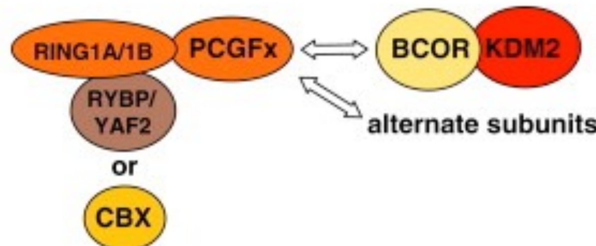
2) Nucleosome dynamics including nucleosome positioning, organization, *compaction*, and turnover/occupancy rates *may be more central to determining transcriptional status*. For example, recent studies on polynucleosome compaction prompt reconsideration of a long-standing idea that PcG silencing could involve altered template accessibility to the transcriptional apparatus and requisite cofactors.

Compositions and Activities of PcG Complexes









(A) The PRC2 family of complexes. Core subunits are in lavender, and arrows depict association of *optional* subunits.

(B–D) The PRC1 family of complexes is depicted.

(**B**) **Canonical** PRC1 from *Drosophila* (left) and *human* (right) with four core subunits including a PC homolog (CBX in mammals).

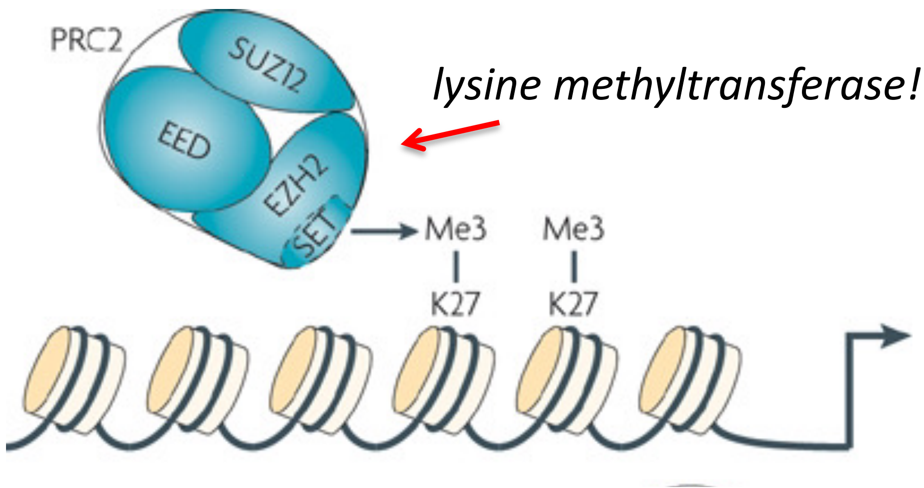
(**C**) Drosophila PRC1 **variant** with KDM2 subunit.

(**D**) Human PRC1 variants that contain KDM2 and/or RYBP subunits. In human PRC1 complexes, assembly of RYBP and CBX subunits are mutually exclusive. Ubiquitylation occurs on H2AK119 in mammals, corresponding to K118 in Drosophila.

Jeffrey A.Simon and Robert E.Kingston, Mol Cell 2013

The PRC2 enzymatic activity

A central function of PRC2 is to methylate histone H3 on K27, with the trimethylated product (H3K27me3) widely viewed as the operative chromatin mark that accompanies PcG silencing.



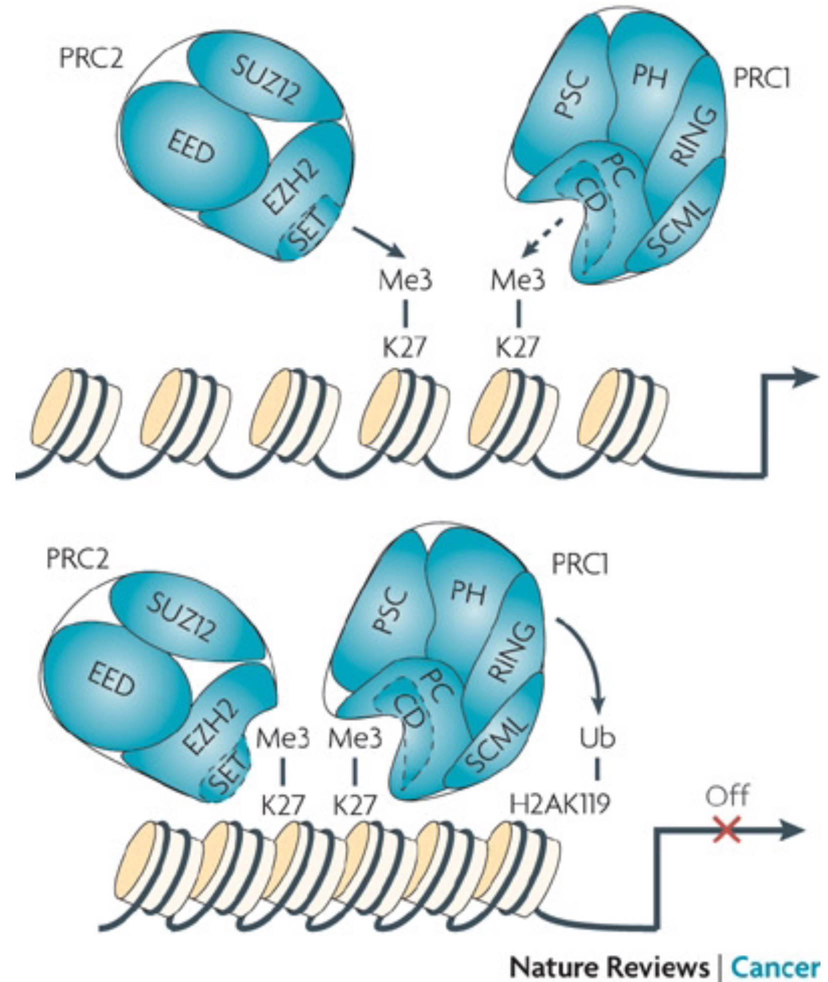
The catalytic subunit, called EZH2 in humans bears a SET domain that houses the enzyme active site. However, the EZH2 subunit is inactive on its own and must be assembled with SUZ12 and EED to produce methyltransferase activity.

Mechanistic Consequences of H3K27me3

We still do not know the mechanisms by which this repressive chromatin mark actually promotes gene repression.

A common view has been that H3K27me3 is instrumental in recruiting PRC1 to chromatin sites.

STILL UNDER DEBATE!



PRC2/K27me3 impedes Pol II recruitment to target gene promoters

How does mammalian PRC2, which *lack* DNA binding capacity, recognize chromatin?

It is still not very clear!

The *paradigm* for targeting PcG complexes emerged from studies on the Drosophila HOX clusters. Targeting to these clusters is accomplished by Polycomb response elements (**PREs**), which bind several sequence-specific DNA binding proteins (reviewed in Ringrose and Paro, 2007; Schuettengruber and Cavalli, 2009) that recruit PRC2 and PRC1. In *Drosophila*, DNA sequences called Polycomb Response Elements (**PRE**) are targets for PcG protein recruitment when inserted at exogenous locus.

These HOX PREs are located **many kilobases from the promoters** they control, are depleted of nucleosomes, and likely form looping interactions with repressed target genes. At HOX genes and genome wide, the most critical PcG recruiter in Drosophila is the zinc finger protein PHO

As well as the tails, other regions of the histone can also be modified: the **central globular** domain of the histones, which together form the core of the nucleosome, also contain a large number of modification sites

The first core modification to be discovered, **H3 lysine 79 methylation**, has also been the most extensively characterized. It plays a fundamental role in the regulation of chromatin structure

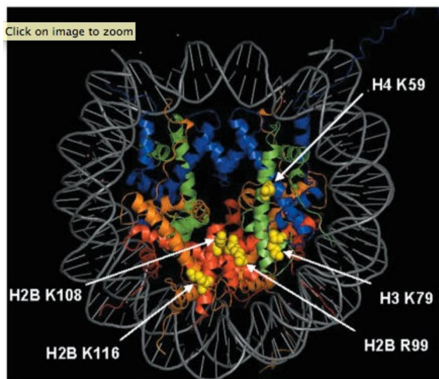
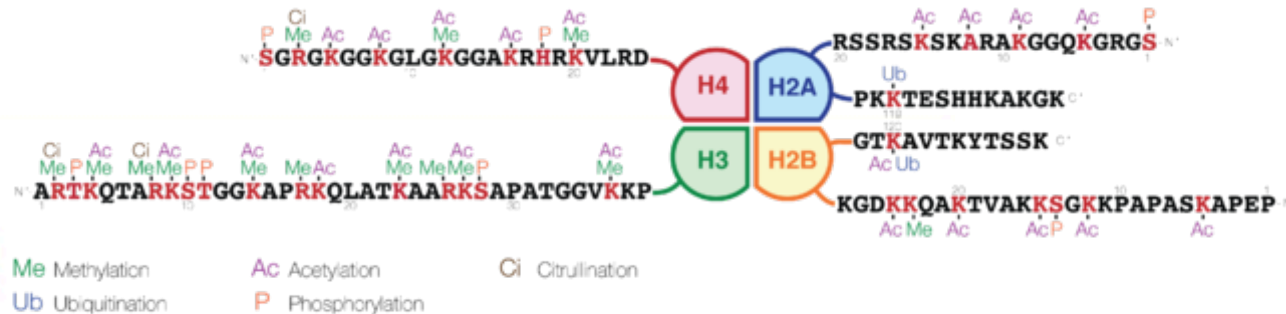


Table 2. Histone Globular Domain Modifications

Histone	Site	Modification	Refs
H2A	H2AK36	Acetylation	[24,96]
	H2AK99	Methylation	[24]
	H2AQ105	Methylation	[89]
	H2AK119	Acetylation	[24]
	H2AK119	Ubiquitylation	[102]
H2B	H2BK40	Methylation	[24]
	H2BK82	Acetylation	[24]
	H2BR96	Methylation	[105]
	H2BK105	Acetylation	[100]
	H2BK113	Acetylation	[100]
	H2BK117	Acetylation	[105]
H3	H3Y41	Phosphorylation	[91]
	H3R42	Methylation	[88]
	H3T45	Phosphorylation	[92]
	H3R53	Methylation	[24]
	H3K56	Acetylation	[64]
	H3K56	Methylation	[56]
	H3K64	Acetylation	[25]
	H3K64	Methylation	[81]
	H3K79	Methylation	[26–28]
	H3K115	Acetylation	[24]
	H3T118	Phosphorylation	[94–96]
	H3K122	Acetylation	[23]
H4	H4K31	Acetylation	[24]
	H4S47	Phosphorylation	[116]
	H4K59	Methylation	[96]
	H4K77	Acetylation	[24]
	H4K79	Acetylation	[117]
	H4K91	Acetylation	[117,118]
	H4R92	Methylation	[24]

The readout of the histone post-transcriptional modifications



Type of modification

- Which amino-acid
- Number of modifications (me)

Position in genome

- Promoter: H3K36me, H3K9me are repressive
- Coding region: H3K36me, H3K9me are activating and prevent cryptic initiation of transcription in ORF

Other histone modifications

- combinatorial (occur together)
- H3K4me + H3K9me: transcriptional activation
- H4K20me + H3K9me: heterochromatin formation
- H3K27me + H3K4me: “bivalent” mark in stem cells

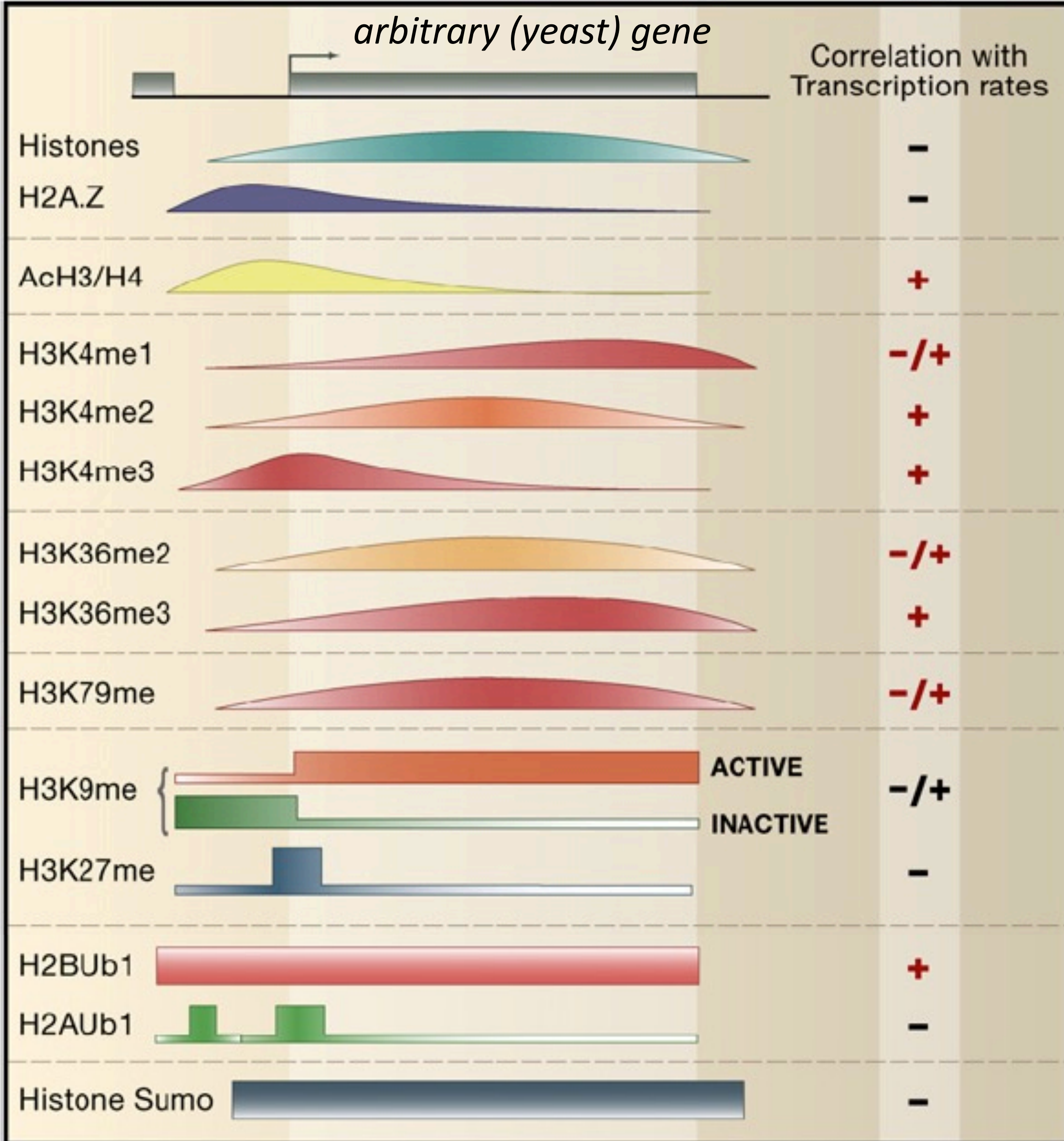
Size of histone modification domain

- large: heritable (can be copied more easily)
  - H3K27me can recruit PRC2 has H3K27me3 activity
  - H3K4me recruits WDR5 (MLL thrithorax): H3K4me

Cycles of modifications

- H2Bub → H2B required for transcriptional elongation

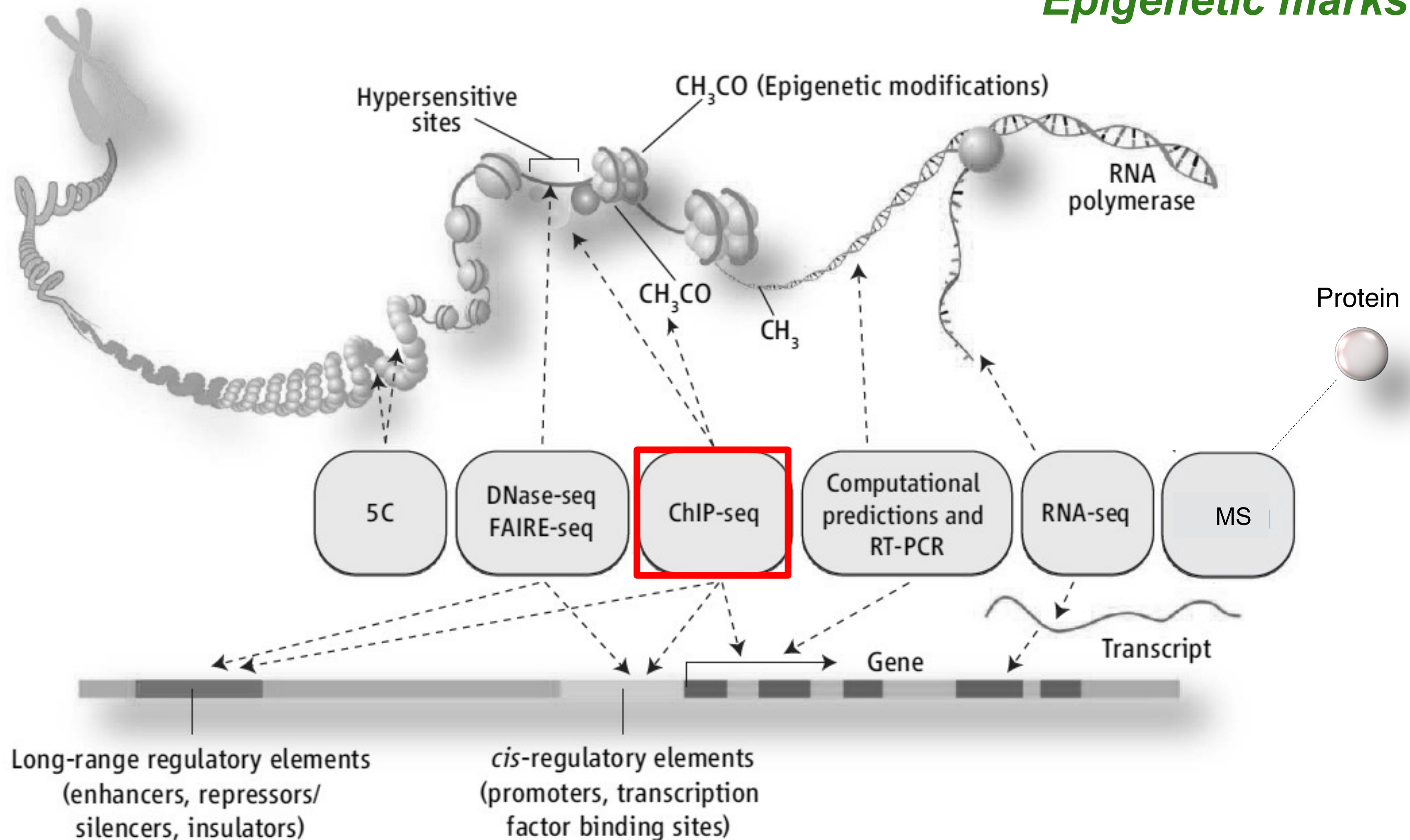
Most modifications are distributed in **distinct localized patterns** within the *upstream region*, the *core promoter, the 5' end* of the open reading frame (ORF) and the *3' end* of the ORF.



The ***location*** of a modification is *tightly* regulated and is crucial for its effect on transcription.

What did we understand from reading the chromatin?

Epigenetic marks



modified from *Science 337:1159-60, 2012*

Main functional classes of metazoan promoters

Based on the configuration of *promoter* signals, transcription start site (*TSS*) *positions*, nucleosome positions and their epigenetic marks, most metazoan promoters can be categorized into three general types:

Promoter type	Dominant gene function
Major promoters	
Type I ('adult')	Tissue-specific expression in adult peripheral tissues
Type II ('ubiquitous')	Broad expression throughout organismal cycle
Type III ('developmentally regulated')	Differentially regulated genes, often regulators in multicellular development and differentiation

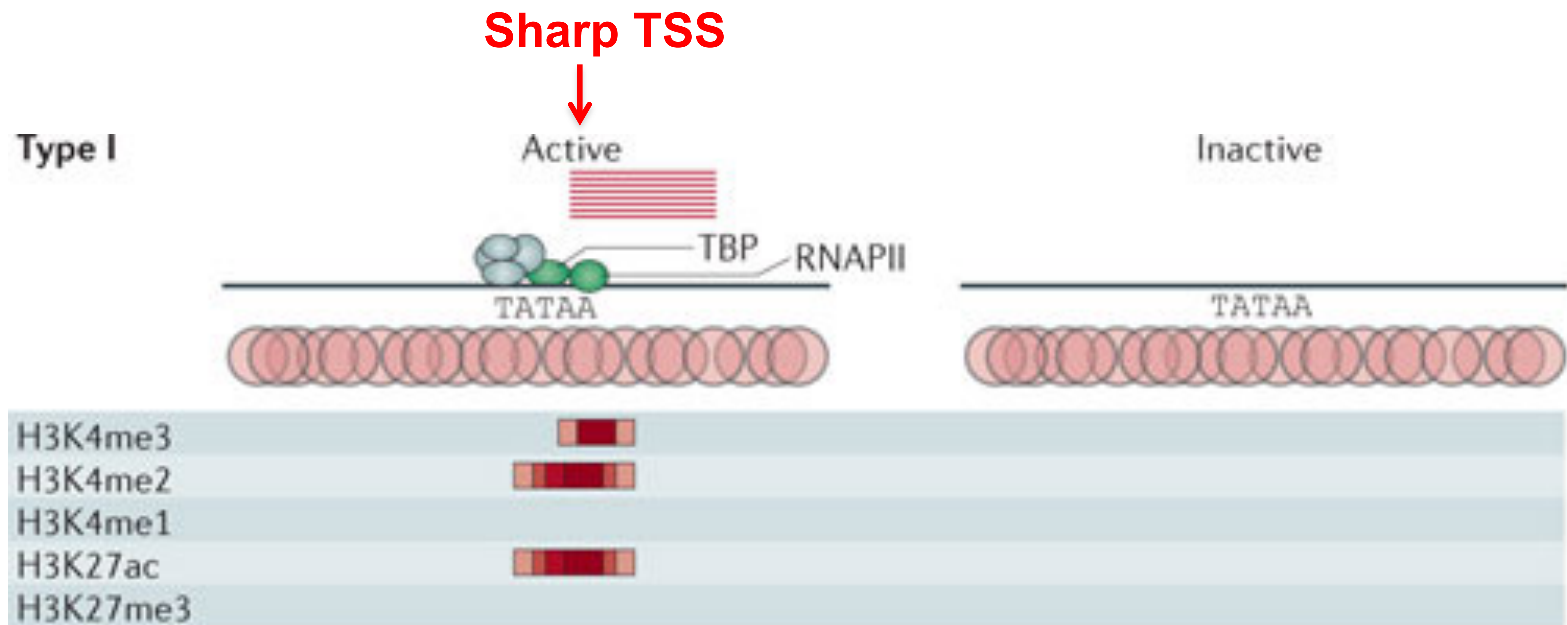
<u>Metazoan promoters: emerging characteristics and insights into transcriptional regulation.</u> Boris Lenhard, Albin Sandelin & Piero Carninci. *Nature Reviews Genetics* **13**, 233-245 (April 2012)

Type I

Promoter type	Dominant gene function	Common properties	Vertebrate-specific
Major promoters			
Type I ('adult')	Tissue-specific expression in adult peripheral tissues	Sharp ('focused') TSS, TATA-box enrichment, disordered nucleosomes	Mostly no CpG islands

Type I ('adult') promoters are most often used for genes that are specifically expressed in terminally differentiated tissues (e.g. liver or skeletal muscle). **H3K4me3** is generally only present downstream of the TSS. There is **no RNAPII** binding at these promoters when the genes are not active. **No CpG islands**

Sharp TSS



Type II

Promoter type	Dominant gene function	Common properties	Vertebrate-specific
Type II ('ubiquitous')	Broad expression throughout organismal cycle	Broad ('dispersed') TSS, ordered nucleosome configuration	CpG islands, TATA-depleted

Type II ('ubiquitous') promoters are most often used for *ubiquitously* expressed genes. They have **H3K4me3** *throughout* their promoters across all tissues. Across classes of genes in vertebrates, **H3K4me3 distribution is almost identical with the span of CpG islands.** Ubiquitously expressed genes generally have *short* CpG islands, and the H3K4me3 mark and CpG island typically only overlap the 5′ end of the gene.

Dispersed TSS

Type II

Active

H3K4me3

H3K4me2

H3K4me1

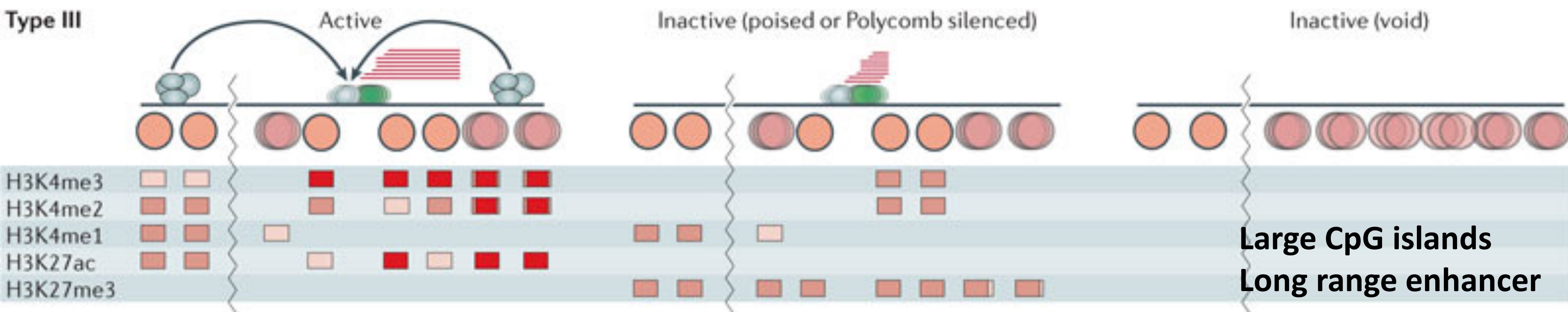
H3K27ac

H3K27me3

Type III

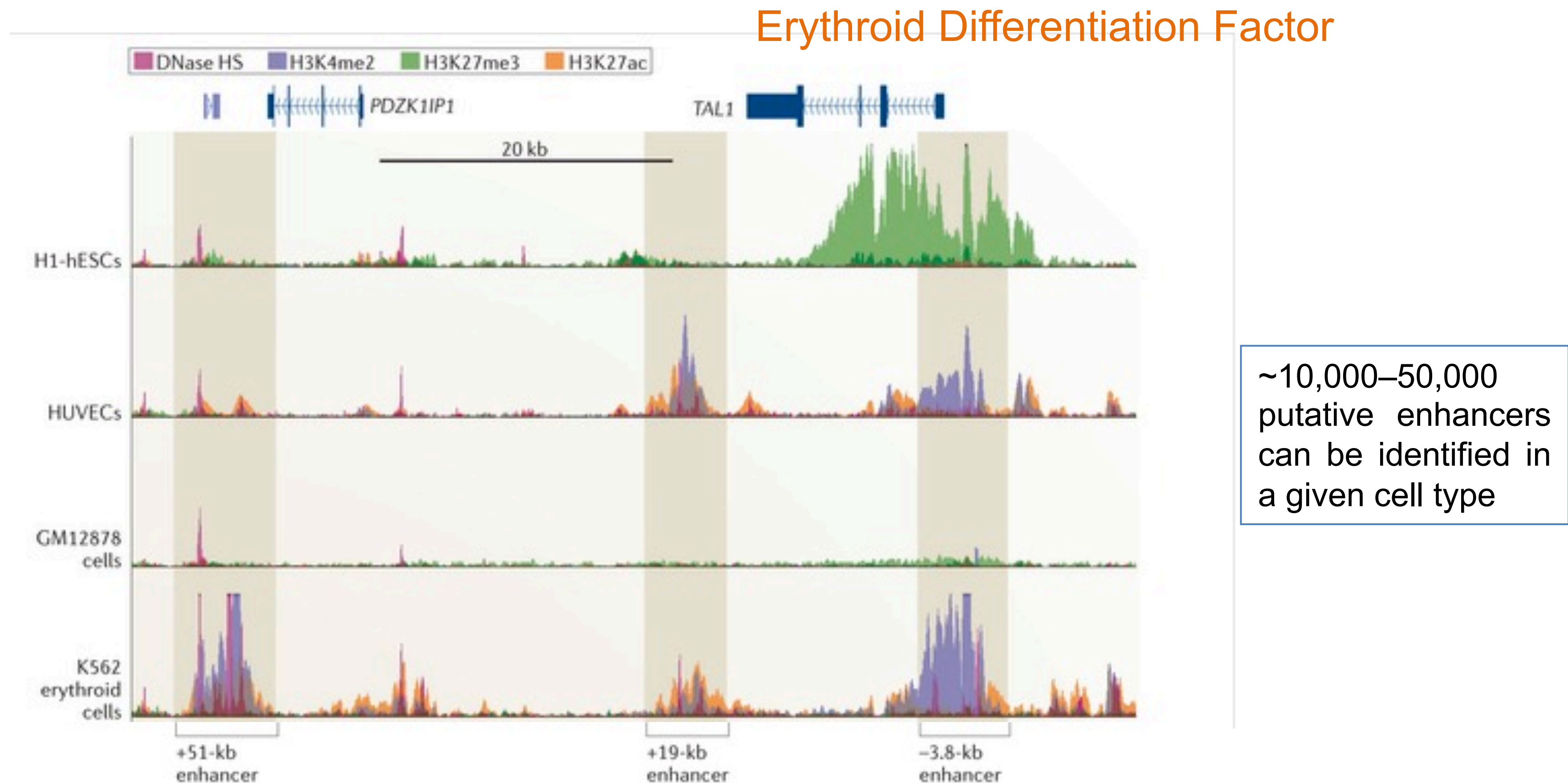
Promoter type	Dominant gene function	Common properties	Vertebrate-specific
Type III ('developmentally regulated')	Differentially regulated genes, often regulators in multicellular development and differentiation	Polycomb repression-regulated genes, broad H3K27me3 marks	Large CpG islands extending into the body of gene

Developmentally regulated genes (with **type III promoters**) have a number of features that are associated with repression by **polycomb group proteins (PcG)**. These features include multiple **large CpG islands**, wide distribution of bound PcG proteins and both H3K27me3 and H3K4me3. Because of the presence of both of these marks, which are associated with repression and activation, respectively, these are described as **bivalent promoters**. The large CpG islands often extend into the body of the genes and are closely tracked by H3K4me3, which is thus not restricted to promoter regions in developmental genes.



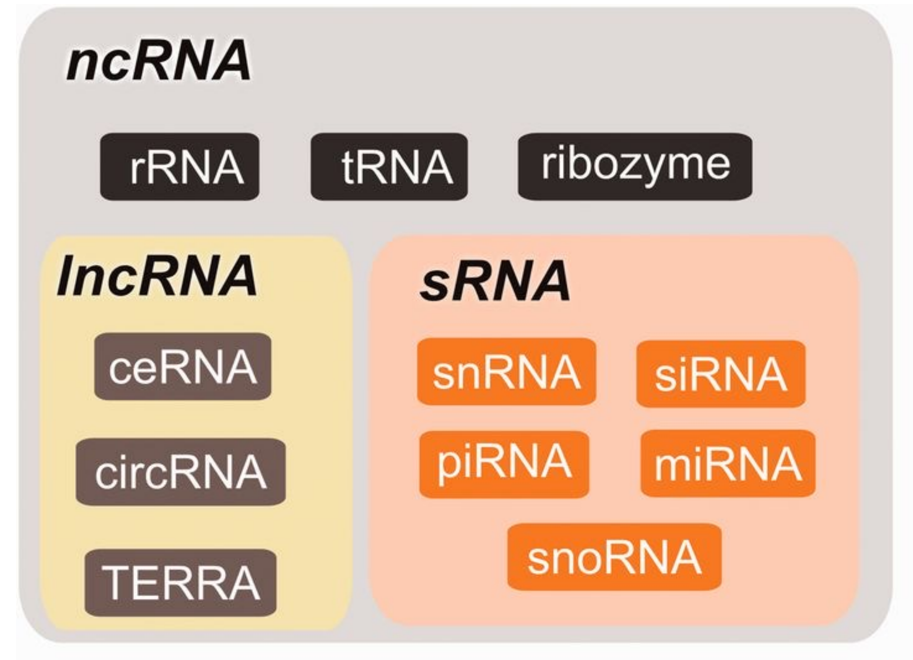
Enhancers

Cell type-specific enhancers are marked by specific **epigenomic** features and chromatin accessibility. In particular, enhancers display enrichment of **H3K4me1** or **H3K4me2** and depletion of H3K4me3 compared with promoters

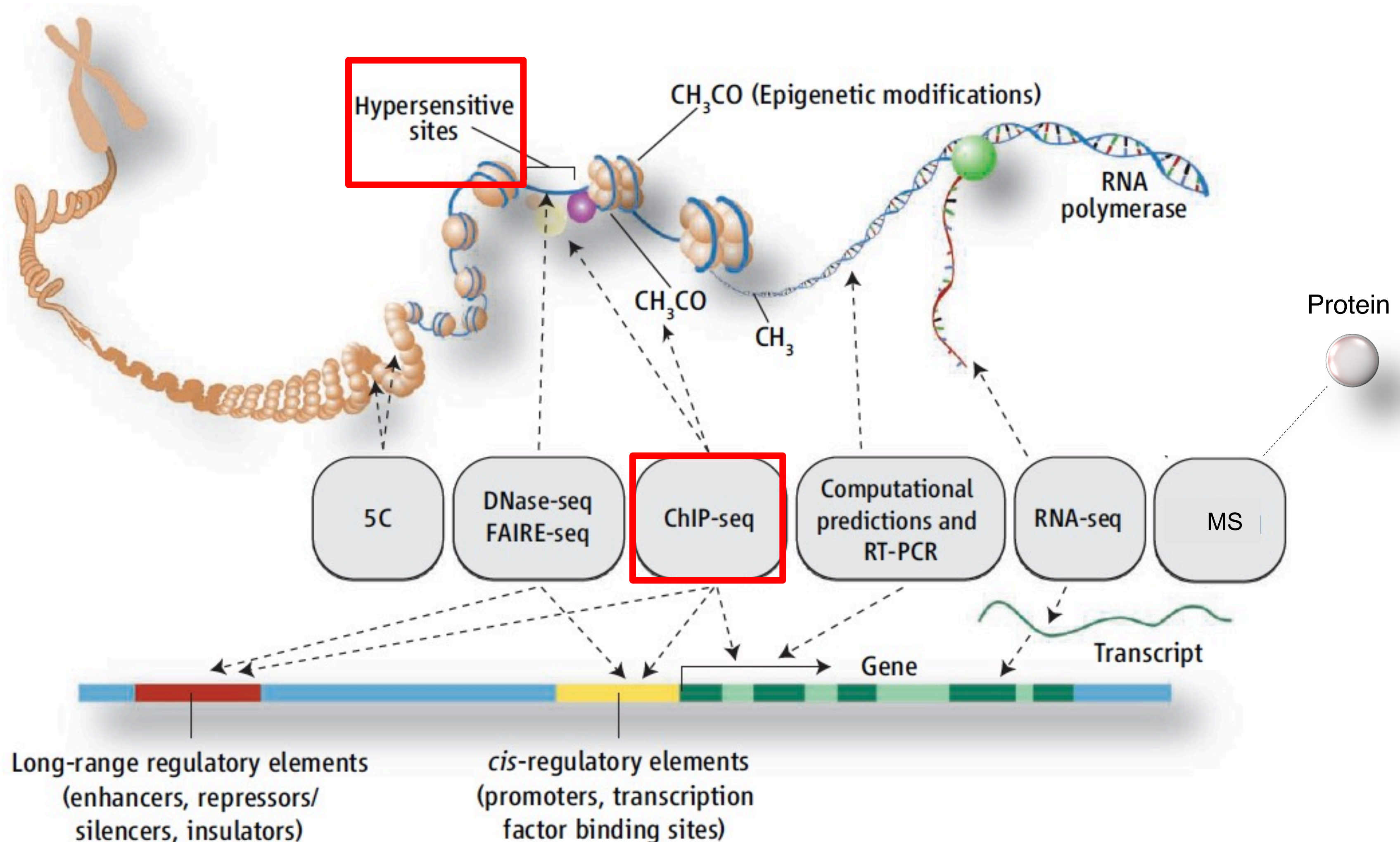


Major highlights:

- Genome is ***pervasively*** transcribed RNA-seq and 5’CAGE
- Genome is organized into ***functional units*** associated with *distinct patterns* of DNA/histone modifications ChIP, DnaseI-seq, 3C
- A large fraction of the ***non coding*** portion is **functional**



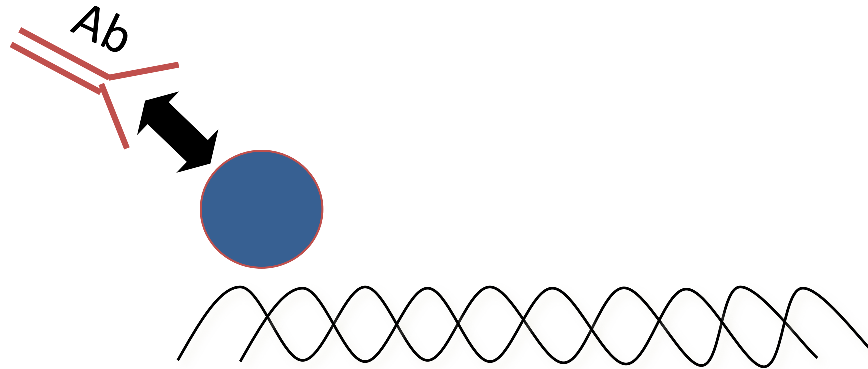
Encyclopedia of DNA Elements (EnCODE) project



modified from *Science 337:1159-60, 2012*

Chromatin ImmunoPrecipitation (ChIP)

A powerful method used to determine ***in vivo*** (*in the chromatin context*) **where** and **when** a particular protein interaction occurs on specific DNA regions

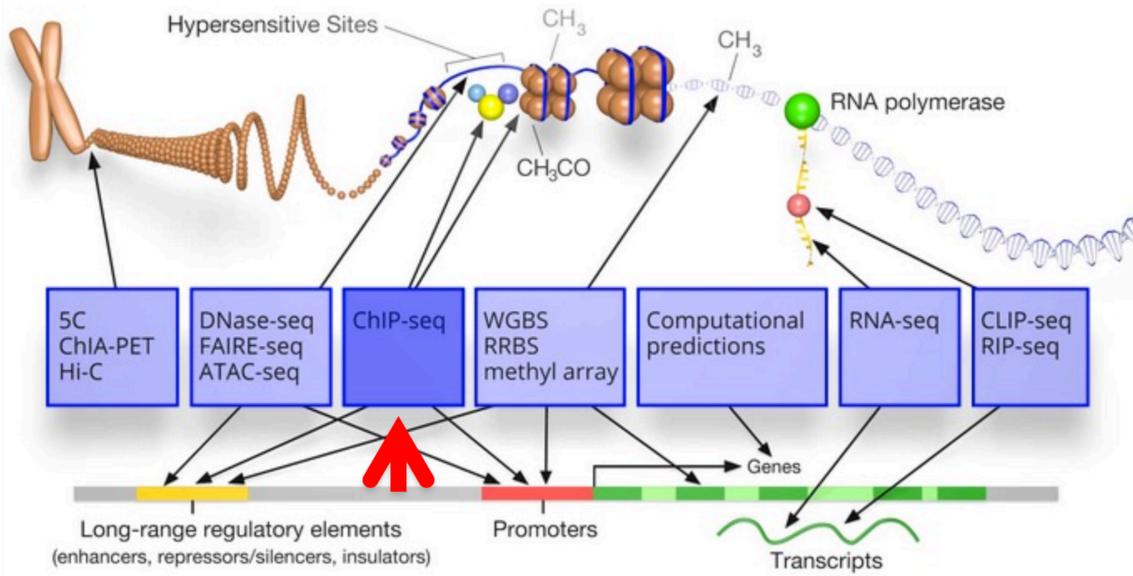


The principle:
selective **enrichment** of a chromatin fraction containing a specific protein

https://www.encodeproject.org/

ENCODE Data Encyclopedia Materials & Methods Help Search...

ENCODE: Encyclopedia of DNA Elements



About ENCODE Project | Getting Started | Experiments

Search ENCODE portal

ENCODE

About ENCODE Encyclopedia | Candidate Regulatory Elements

Search for Candidate Regulatory Elements

Hosted by SCREEN

Human hg19 | Mouse mm10

ChIP sequencing

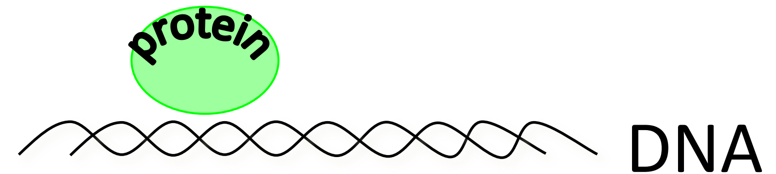
By combining **chromatin immunoprecipitation (ChIP)** assays with **DNA sequencing**, ChIP sequencing (ChIP-Seq) is a powerful method for identifying genome-wide DNA binding sites for transcription factors and other proteins.

Advantages of ChIP-Seq

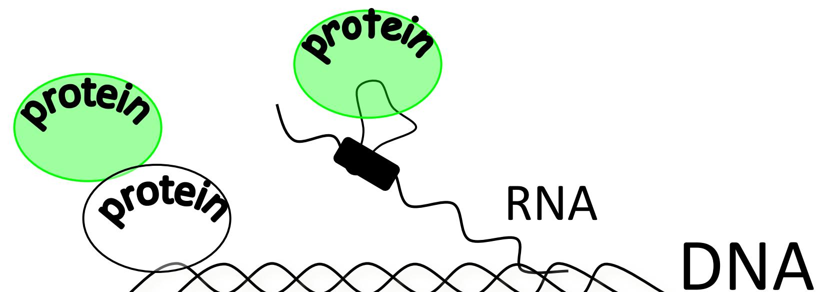
- Captures DNA targets for *transcription factors* or *histone modifications* across the entire genome of any organism
- Defines transcription factor binding sites
- Reveals gene regulatory networks in combination with RNA sequencing and methylation analysis

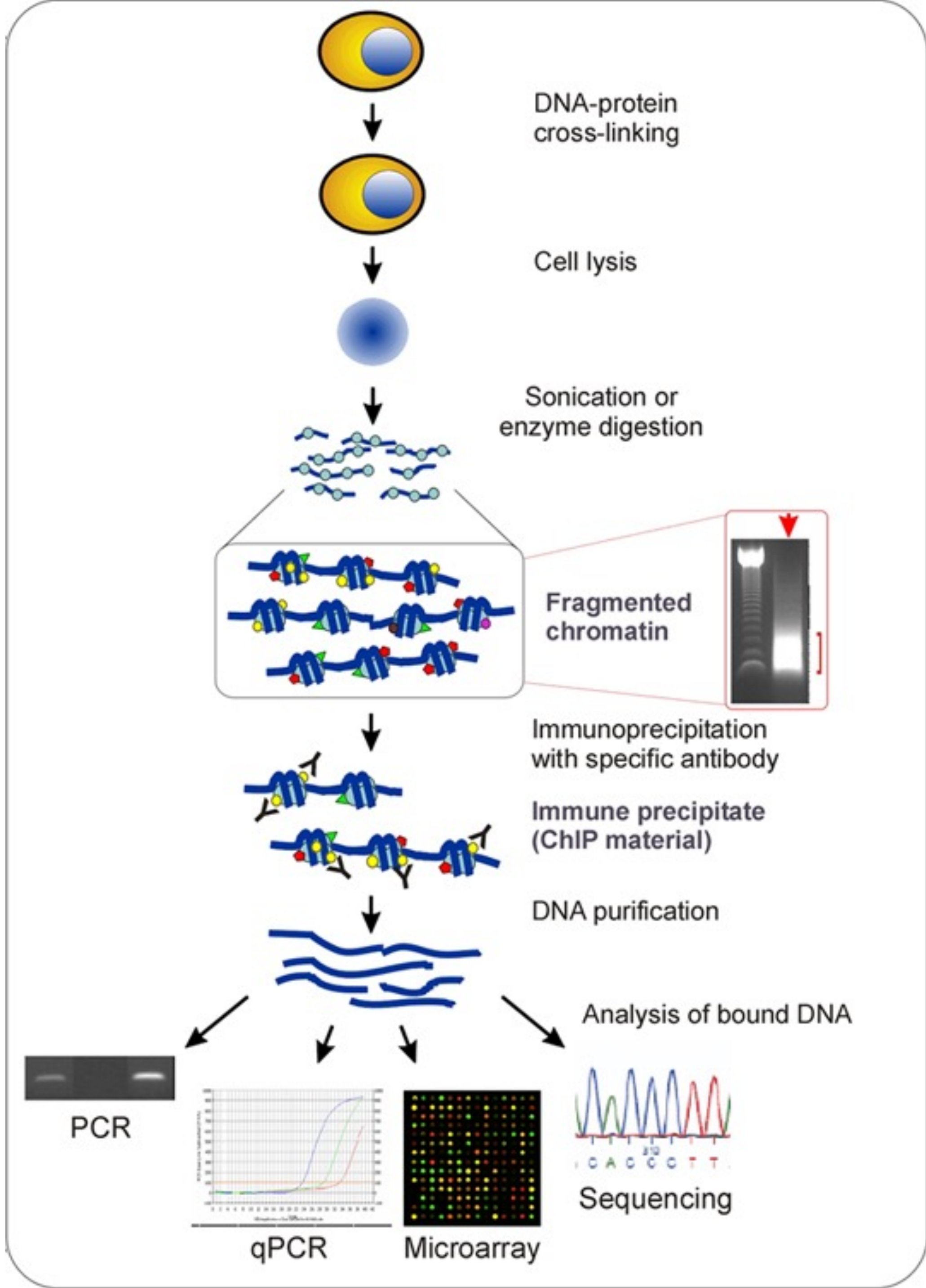
*x*ChIP is used to map:

1. Proteins **directly** bound to DNA
   (i.e. histone modifications, transcription/replication factors)



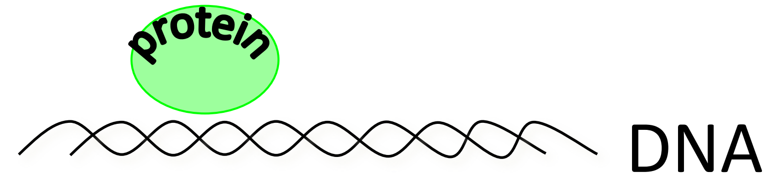
2. Proteins **not directly** bound to DNA
   (i.e. mRNA processing factors).



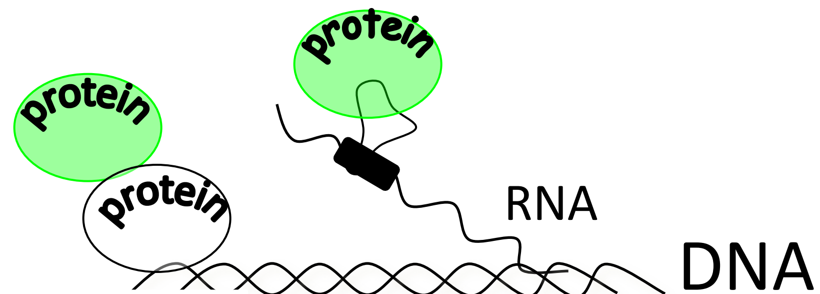

DNA-protein
cross-linking
Cell lysis
Sonication or
enzyme digestion
Fragmented
chromatin
Immunoprecipitation
with specific antibody
Immune precipitate
(ChIP material)
DNA purification
Analysis of bound DNA
PCR
qPCR
Microarray
Sequencing

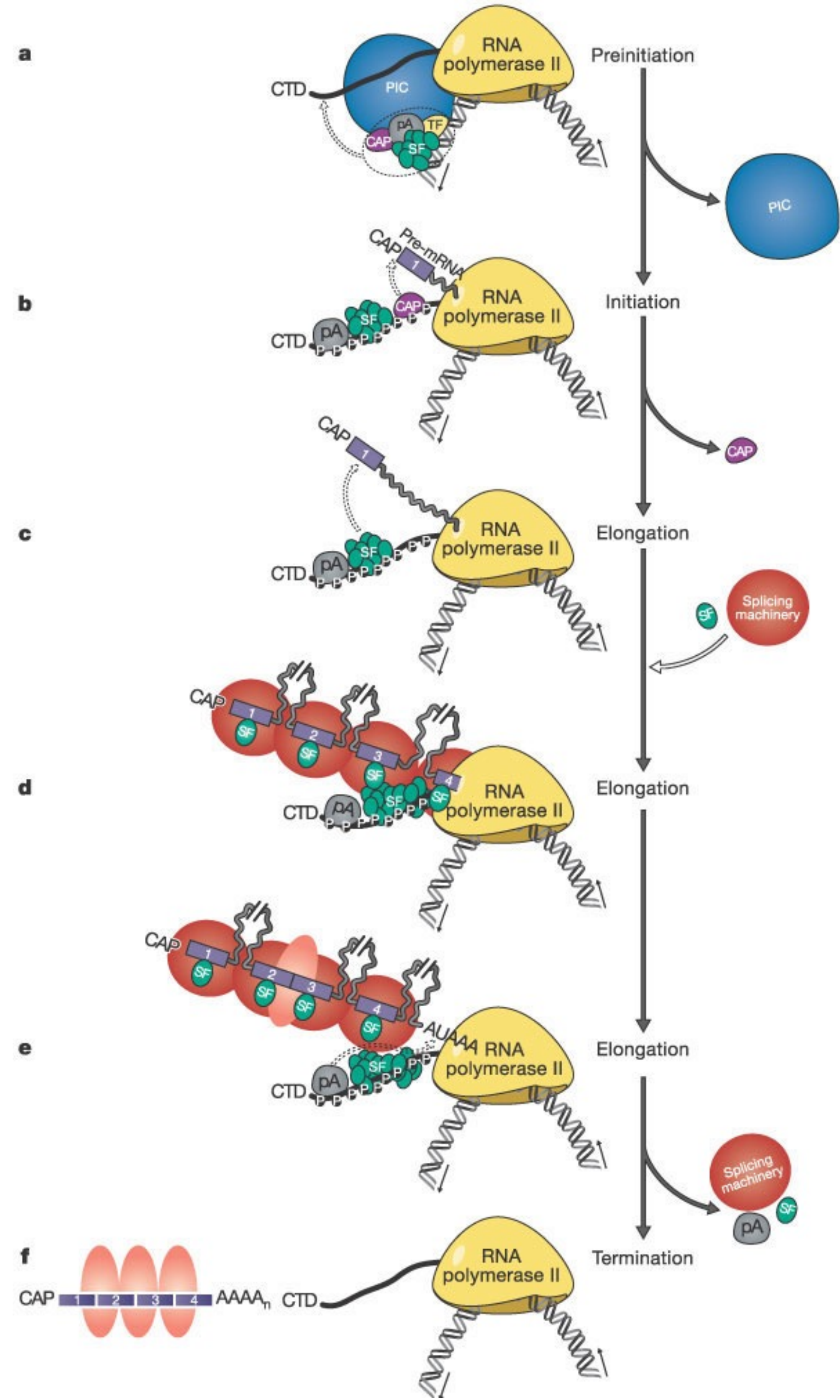
X-ChIP is used to map:

1. Proteins **directly** bound to DNA
   (i.e. transcription, replication, modification)



2. Proteins **not directly** bound to DNA
   (i.e. mRNA processing factors).





During and after transcription, RNAs undergo to **multiple processing** steps which are coordinated by RNA-binding proteins.

Coupling plays a critical role in gene expression dramatically increasing the specificity of enzymatic reactions.

A key task is to ***map*** **co-transcriptional interactions**

Tom Maniatis and Robin Reed
Nature **416**, 499-506 (4 April 2002)

WORK FLOW

crosslinking with formaldehyde (1% final)

lysis, chromatin extraction

chromatin shearing

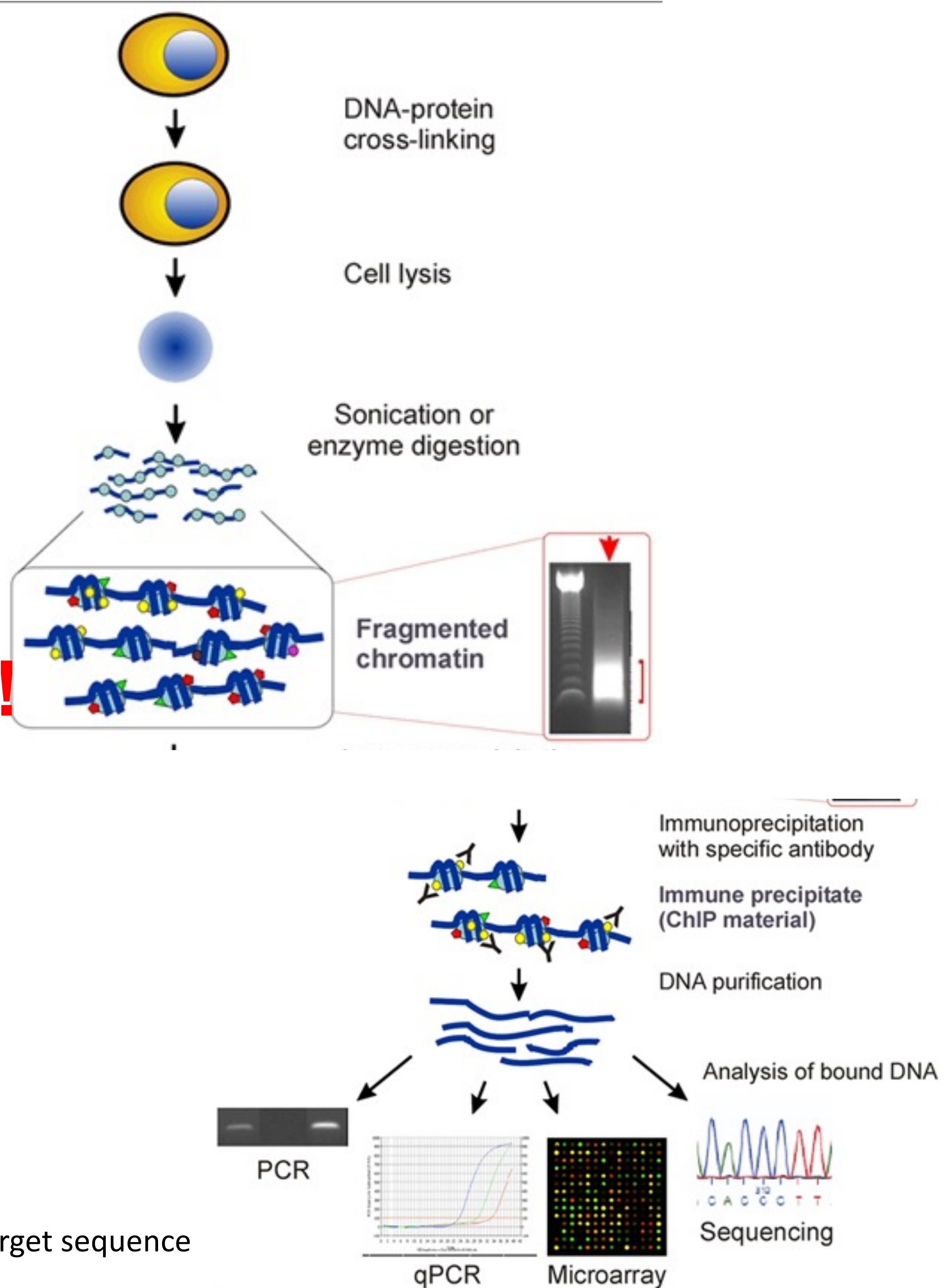
RNAse treatment!!!!

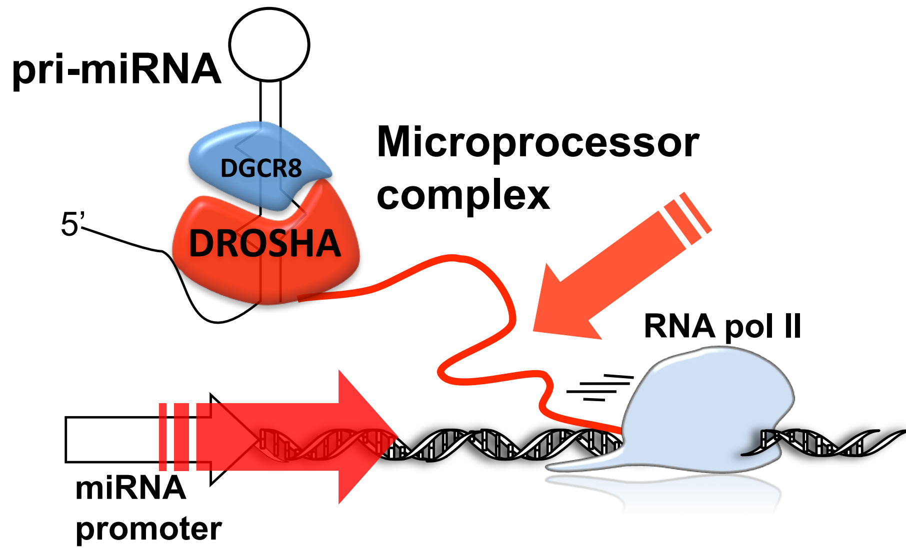
immunoprecipitation of resulting chromatin

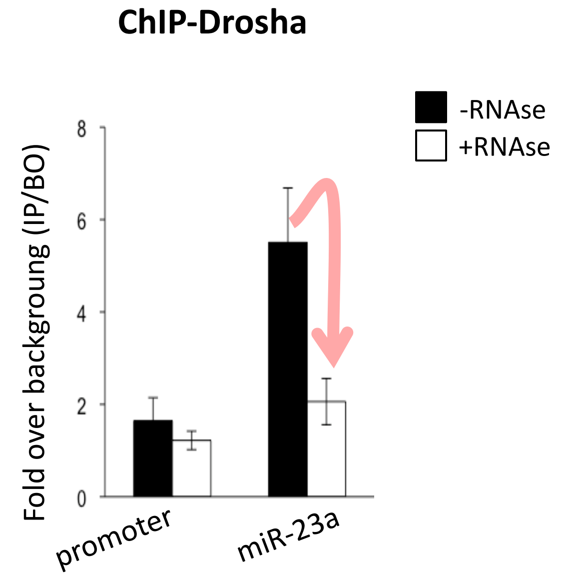
reverse cross-linking (heat treatment)

proteins digestion (Proteinase K) and DNA isolation

quantification to determine the *enrichment* of DNA target sequence



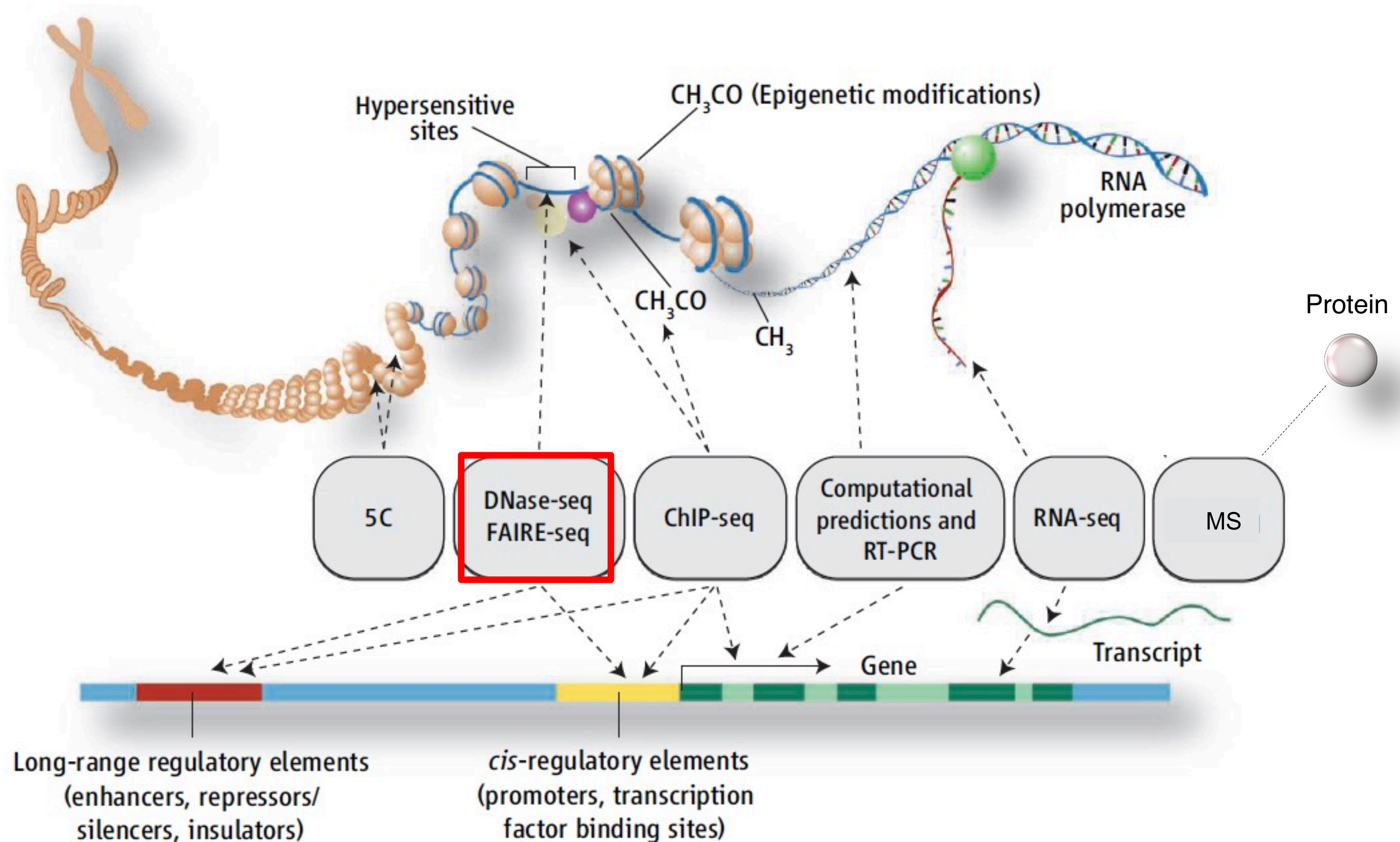




- **Drosha is co-transcriptionally recruited on miR loci**
- **miR-promoter favours Drosha recruitment to chromatin**
- **Drosha recruitment is RNA-dependent**

Ballarino M, Pagano F, et al **2009** *Mol Cell Biol*

Encyclopedia of DNA Elements (EnCODE) project

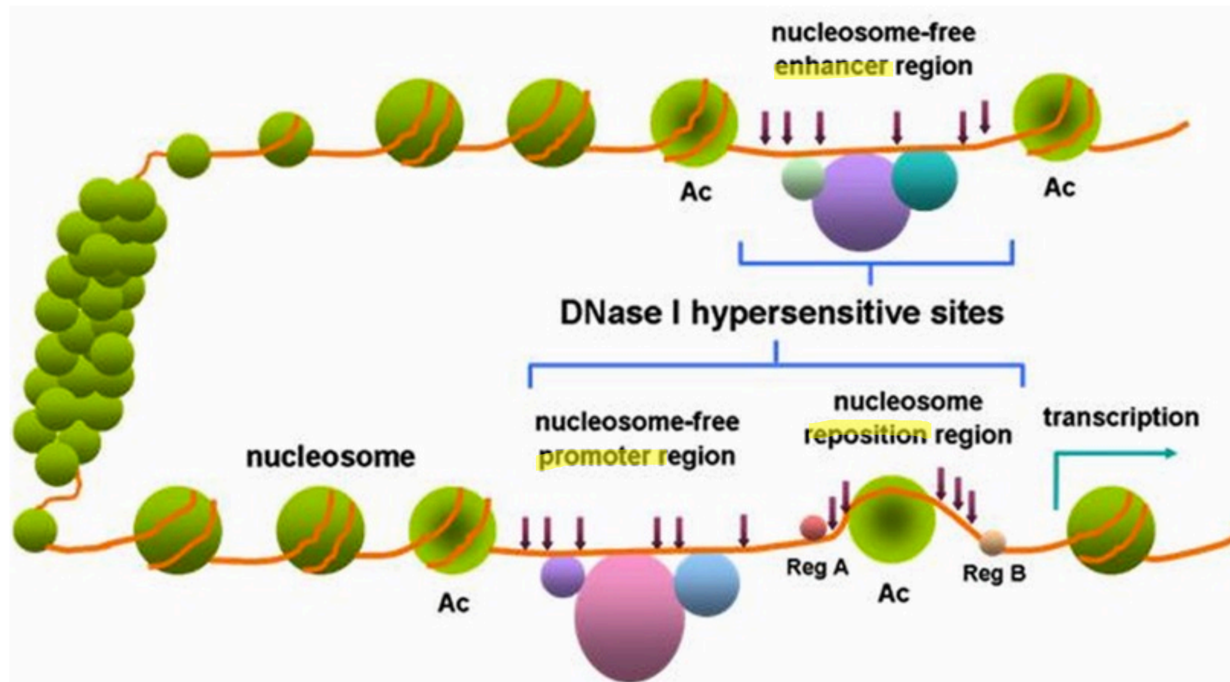


modified from *Science 337:1159-60, 2012*

Detecting cis-regulatory elements by their sensitivity to DNase I

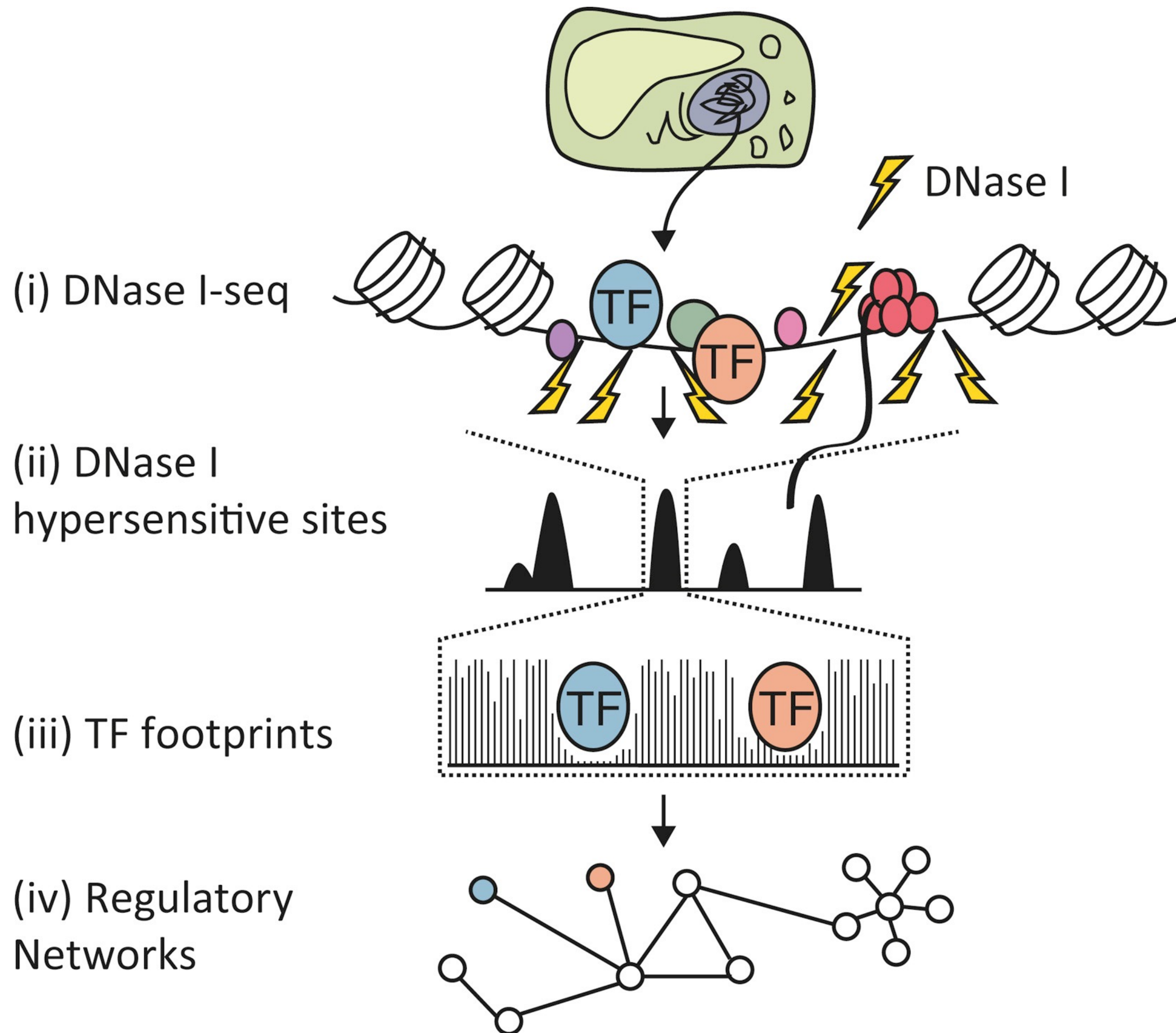
Active ***cis-regulatory*** elements, including promoters, enhancers, insulators, silencers, and locus control regions, can be detected by their characteristic *hypersensitivity* to the endonuclease **DNase I**. **DNase I hypersensitive sites** (DHSs) are regions of chromatin that are *sensitive* to cleavage by the **DNase I** enzyme.

In these specific regions of the genome, chromatin has ***lost*** its *condensed* structure, exposing the DNA and making it accessible. This raises the availability of DNA to degradation by enzymes, such as DNase I.

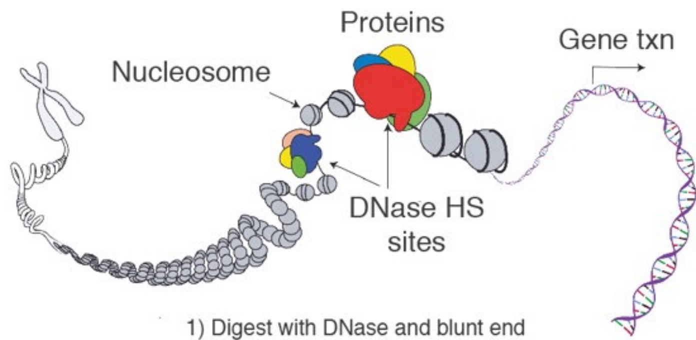


DHSs are functionally related to transcriptional activity, since this remodeled state is necessary for the binding of proteins such as transcription factors.

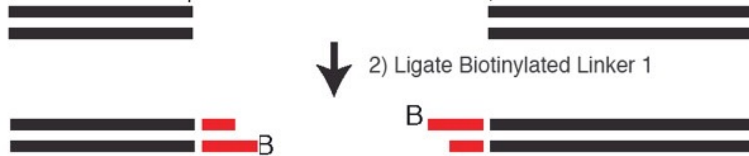
Schematic of DNase I-seq-derived data



(i) Nuclei are harvested from plant tissues and treated with the endonuclease DNase I. (ii) Regulatory regions are hypersensitive to cleavage by DNase I. (iii) Protein-bound regions within DNase I hypersensitive sites are protected from DNase I cleavage leaving detectable “footprints.” (iv) Footprint and TF motif information can be integrated to generate TF-to-TF regulatory networks.



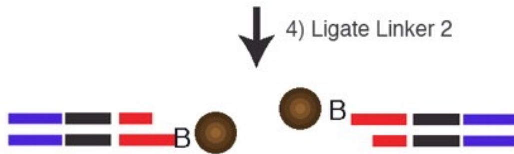
1) Digest with DNase and blunt end



2) Ligate Biotinylated Linker 1

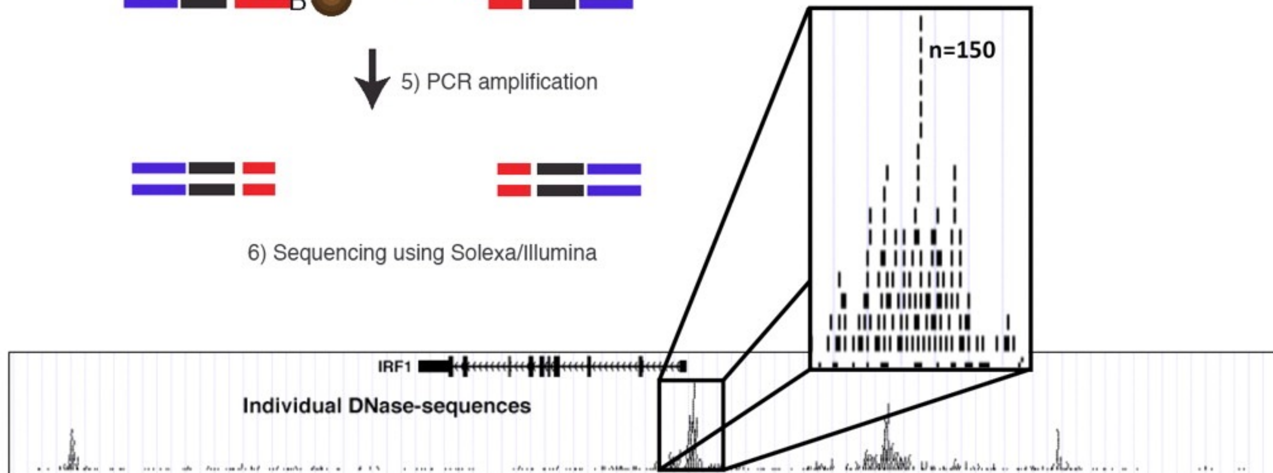
3) MmeI digested, bind to Dynal beads

4) Ligate Linker 2



5) PCR amplification

6) Sequencing using Solexa/Illumina

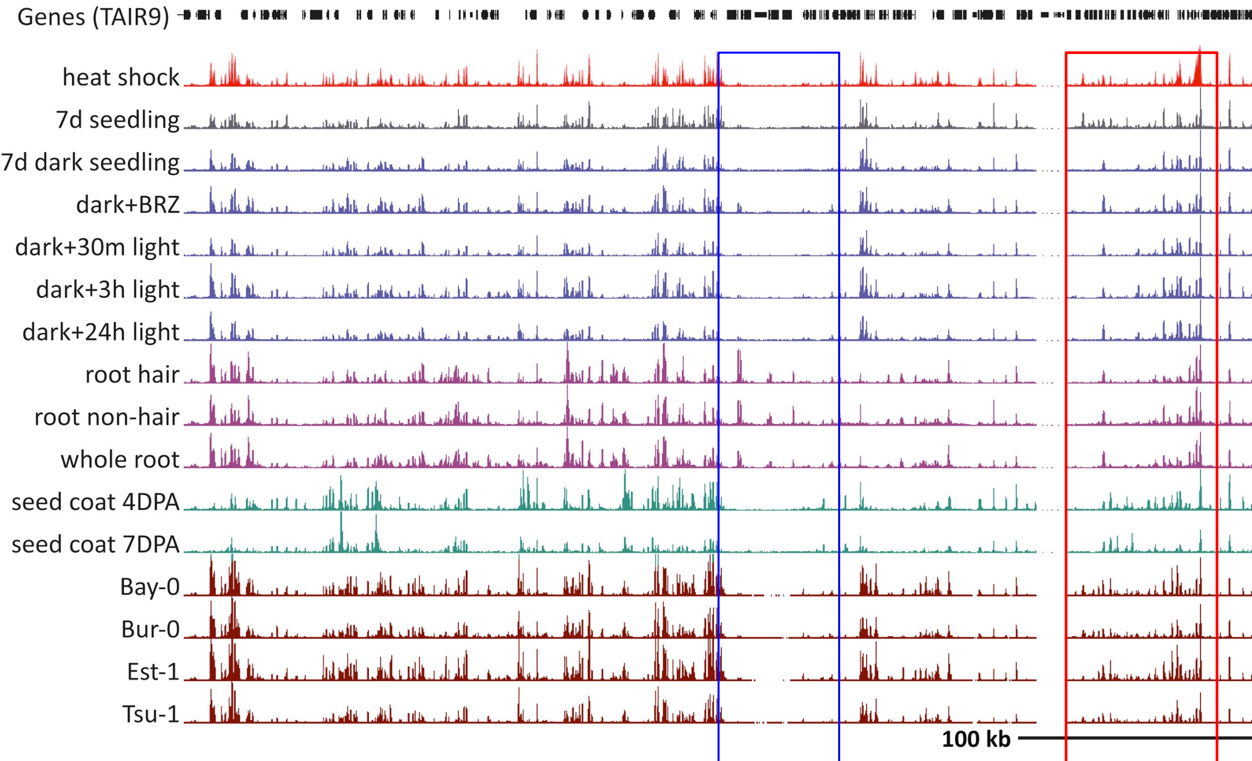


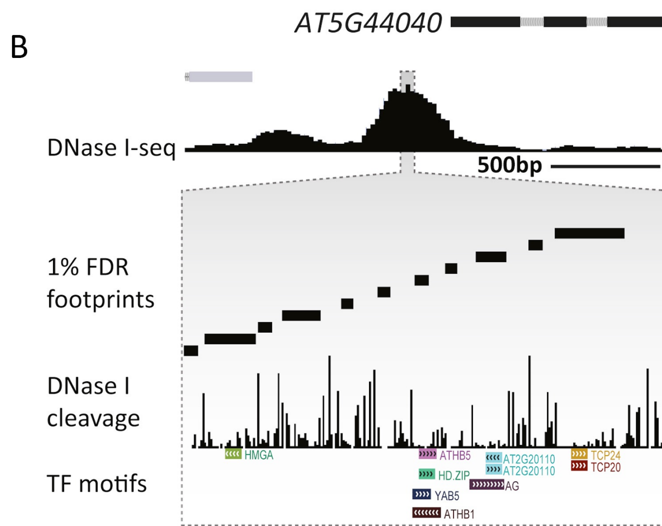
Article ALERT!

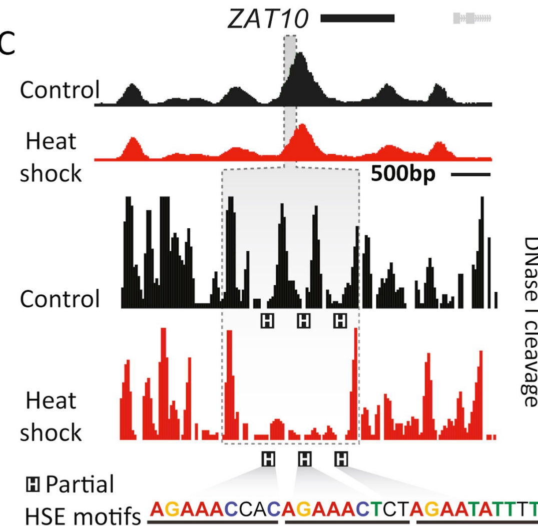
DNase-seq: A High-Resolution Technique for Mapping Active Gene

Regulatory Elements across the Genome from Mammalian Cells
Lingyun Song and Gregory E. Crawford. Cold Spring Harb Protoc; 2010;
doi:10.1101/pdb.prot5384

DNaseI-seq is a valuable orthogonal method to facilitate TF motif discovery.

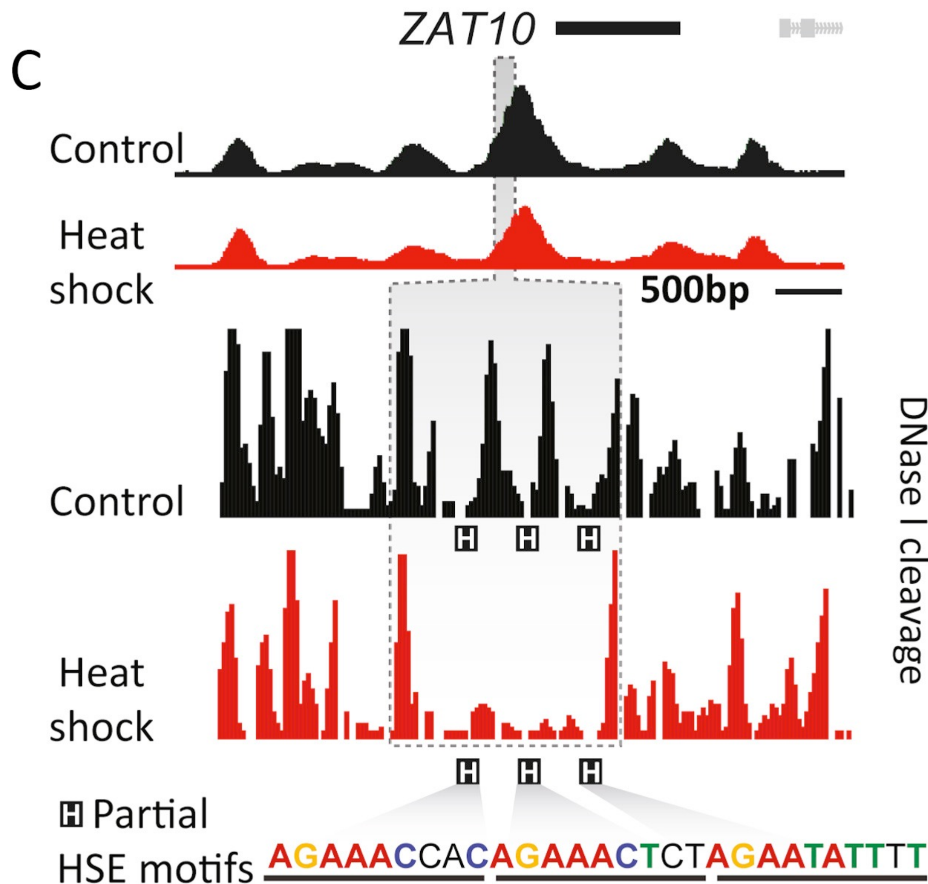






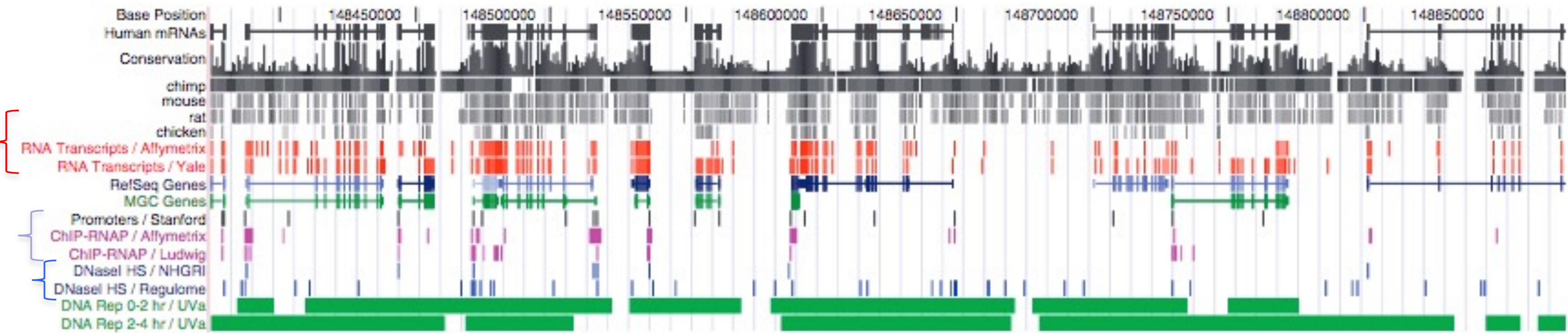
(A) Browser view of a large genomic region of DNase I-seq data in seedlings, tissues, cell types, and *A. thaliana* accessions taken from www.plantregulome.org. This region displays **different patterns of chromatin accessibility** across the sample types. Many DNase I hypersensitive regions **are also *shared*** among samples. (B) A DNase I hypersensitive site in 7 day old light grown seedling in the promoter region of AT5G44040, a gene of unknown function. Footprints have been systematically detected and TF motifs from TRANSFAC [83] and a recent protein binding microarray study [39] are shown. (C) **A differentially accessible region in the promoter of ZAT10 during heat shock (red).** Partial heat shock elements (HSEs) are observed within the differentially accessible region. Control sample is shown in black.

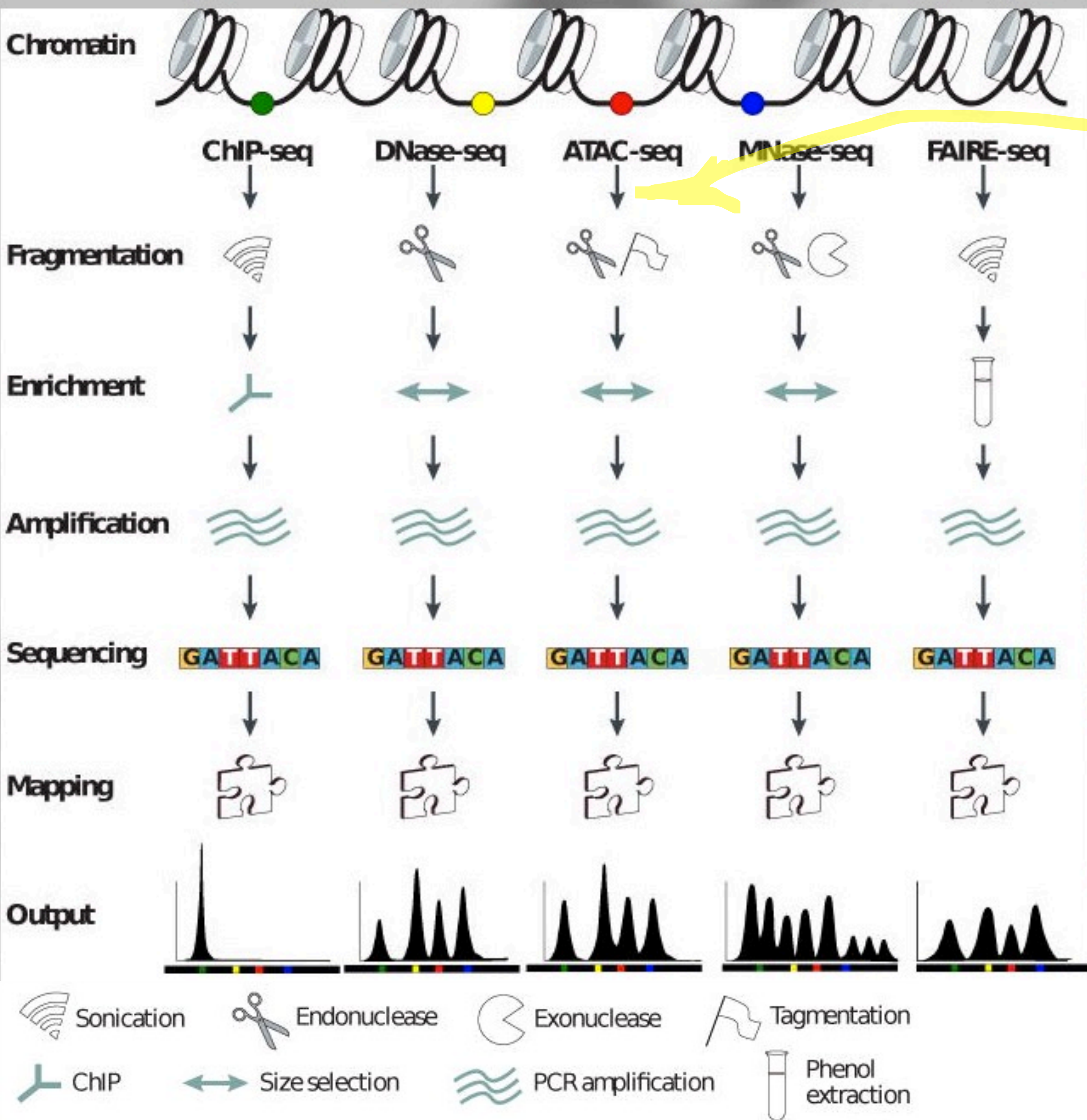
DNaseI-seq is a valuable orthogonal method to facilitate TF motif discovery.



A differentially accessible region in the promoter of ZAT10 during heat shock (red). Partial heat shock elements (HSEs) are observed within the differentially accessible region. Control sample is shown in black.

UCSC Genome Browser display of representative ENCODE data





ATAC-seq employs the mutated hyperactive transposase. During "tagmentation", Tn5 transposase cleaves and tags double-stranded DNA with sequencing adaptors