Chromatin, histone modifications, and chromatin accessibility

Bio 5488

Michael Meers

2/4/2026

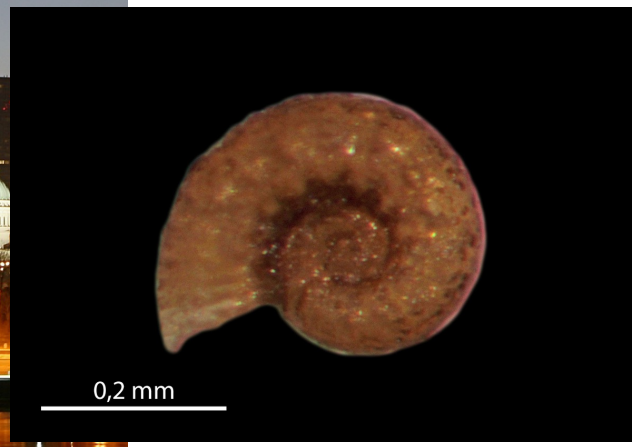
Chromatin: DNA-protein complex that condenses chromosomes

All DNA in cell: 5-6 meters

500,000x condensation

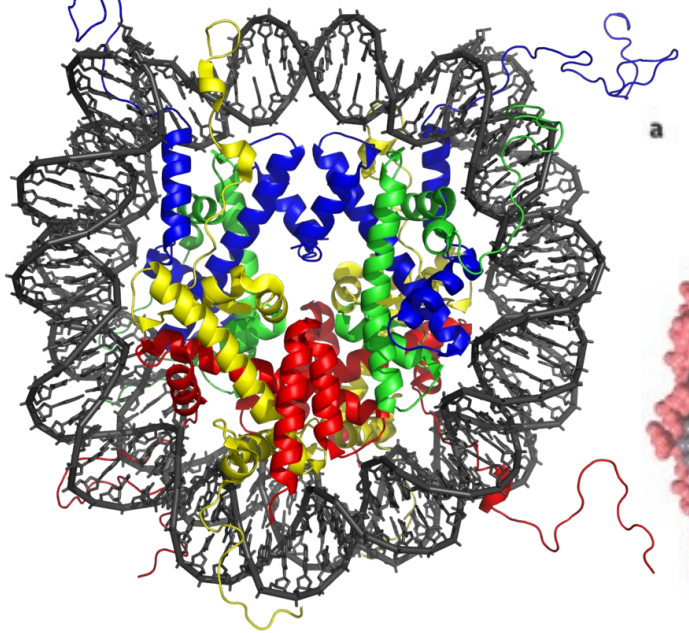
Nucleus: 10 microns

By St_Louis_night_expblend.jpg: Daniel Schwen derivative work: ←fetchcomms - St_Louis_night_expblend.jpg, CC BY-SA 3.0

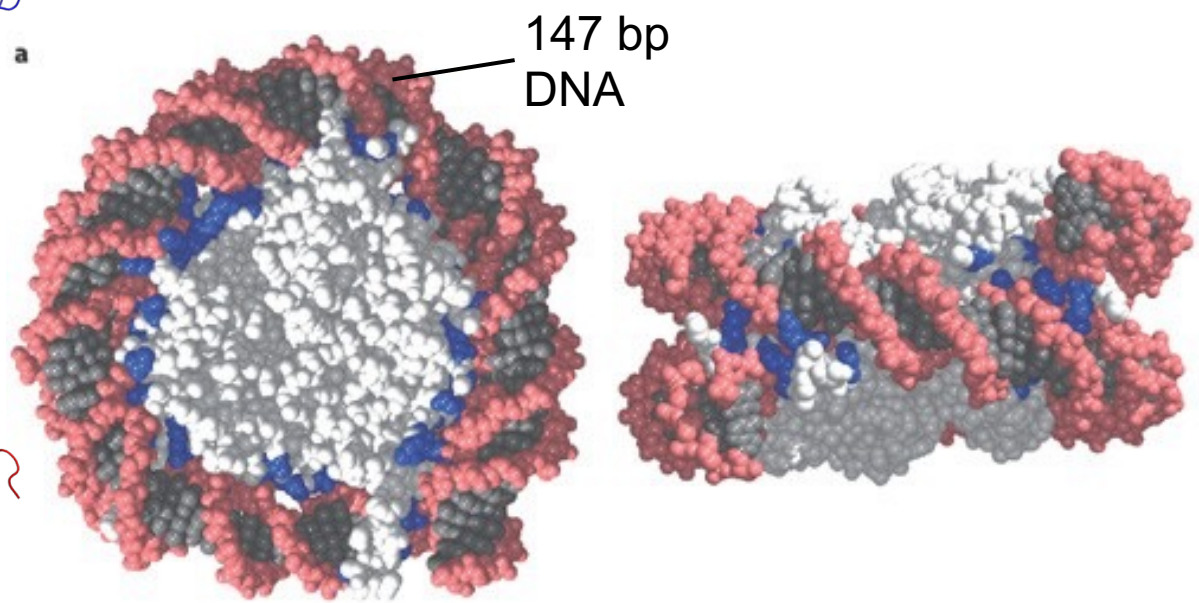


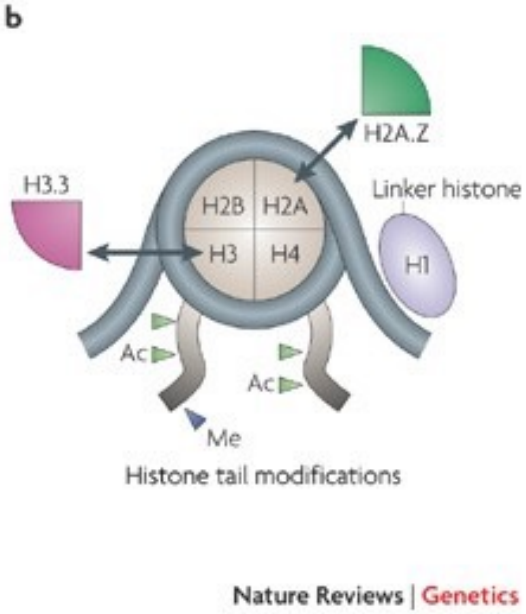
By Marie HENNION - MNHN - Museum national d'Histoire naturelle (2020). The molluscs collection (IM) of the Muséum national d'Histoire naturelle (MNHN - Paris). Version 70.159.

The nucleosome: elemental subunit of chromatin



Luger et al. (1997)

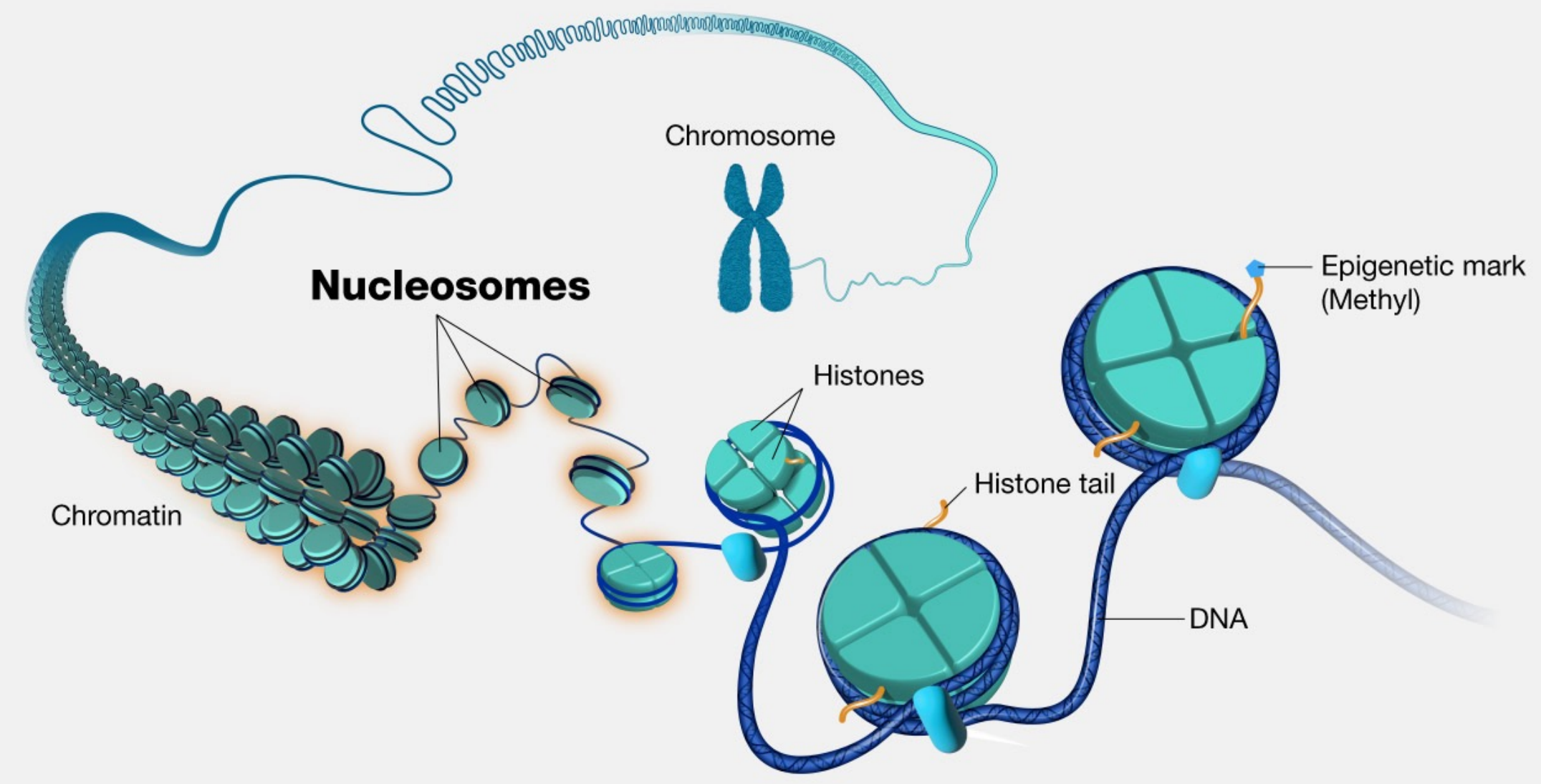




Jiang and Pugh (Nat. Rev. Genet. 2009)

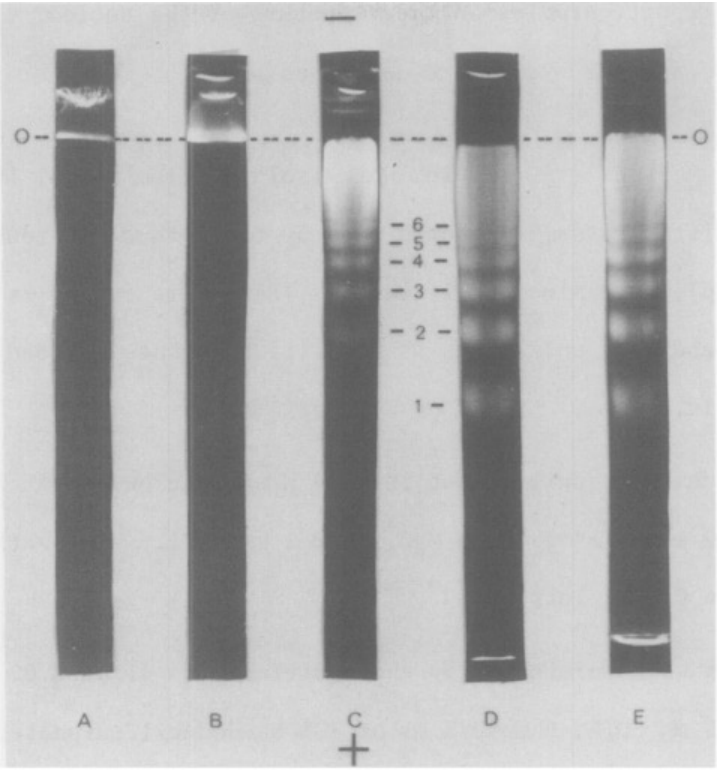
6×10^9 (bp/genome)/
~180 (bp/nucleosome)
= ~ 3.33×10^7 nucleosomes/genome

Chromatin is the protein-DNA complex that accomplishes packaging of DNA in the nucleus

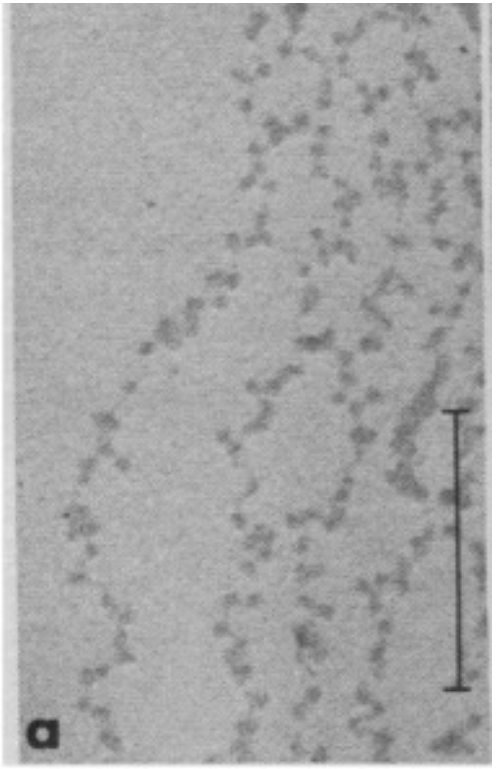


Chromatin Structure: A History

- Repeated chromatin structure *in situ* (Hewish and Burgoyne 1973)
- “Beads on a string” (Olins and Olins 1974)
- Tetramers of H3-H4 and H2A-H2B compact DNA 5-10 fold (Kornberg 1974)



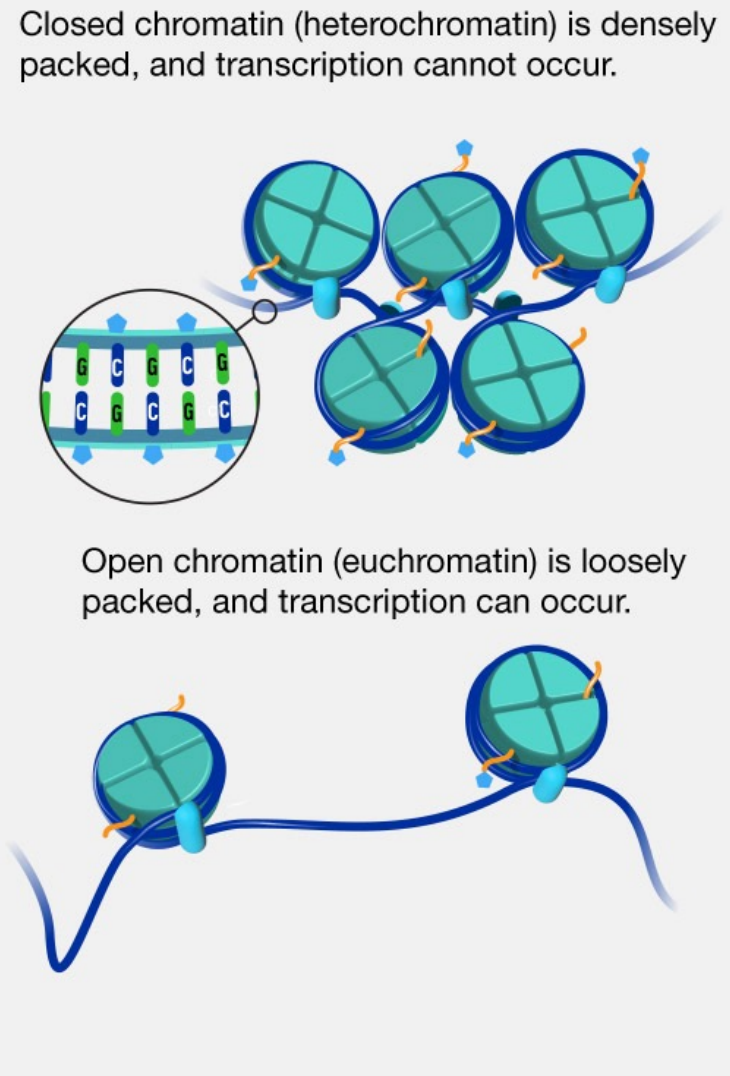
Hewish and Burgoyne (1973)



Olins and Olins (1974)

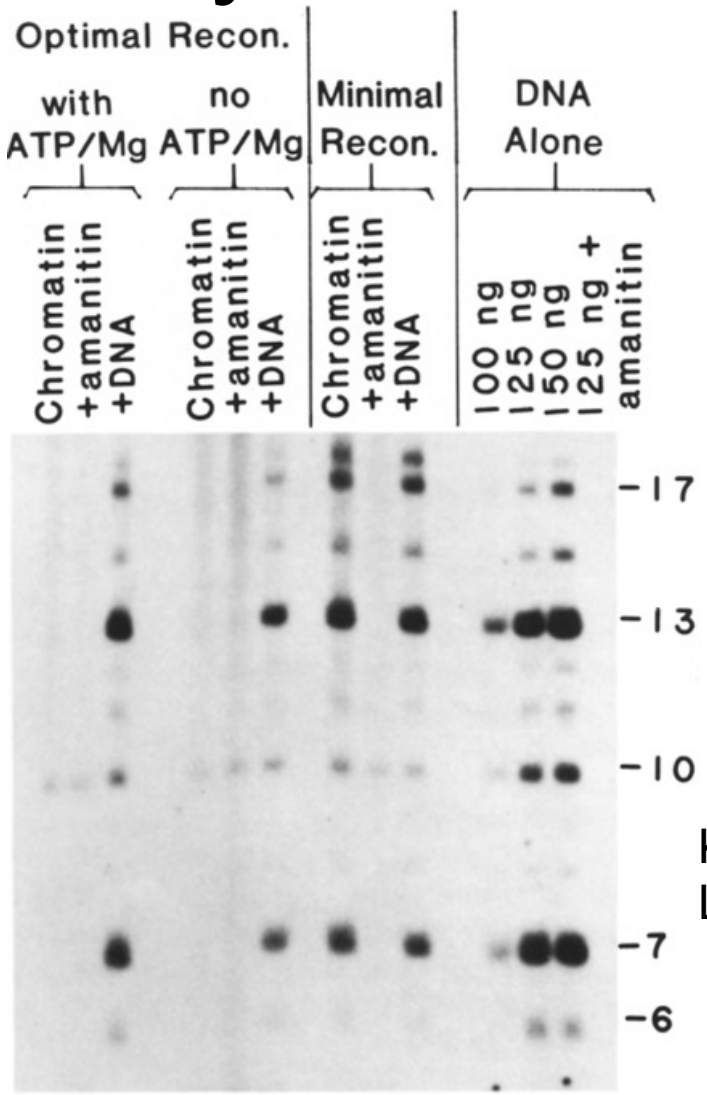
Chromatin Structure: A History

- Globin DNA is differentially nuclease-accessible in different cell types (Weintraub and Groudine 1976)
- Active nuclear transcription complexes are mutually exclusive with histones (Roeder 1982)



Chromatin Structure: A History

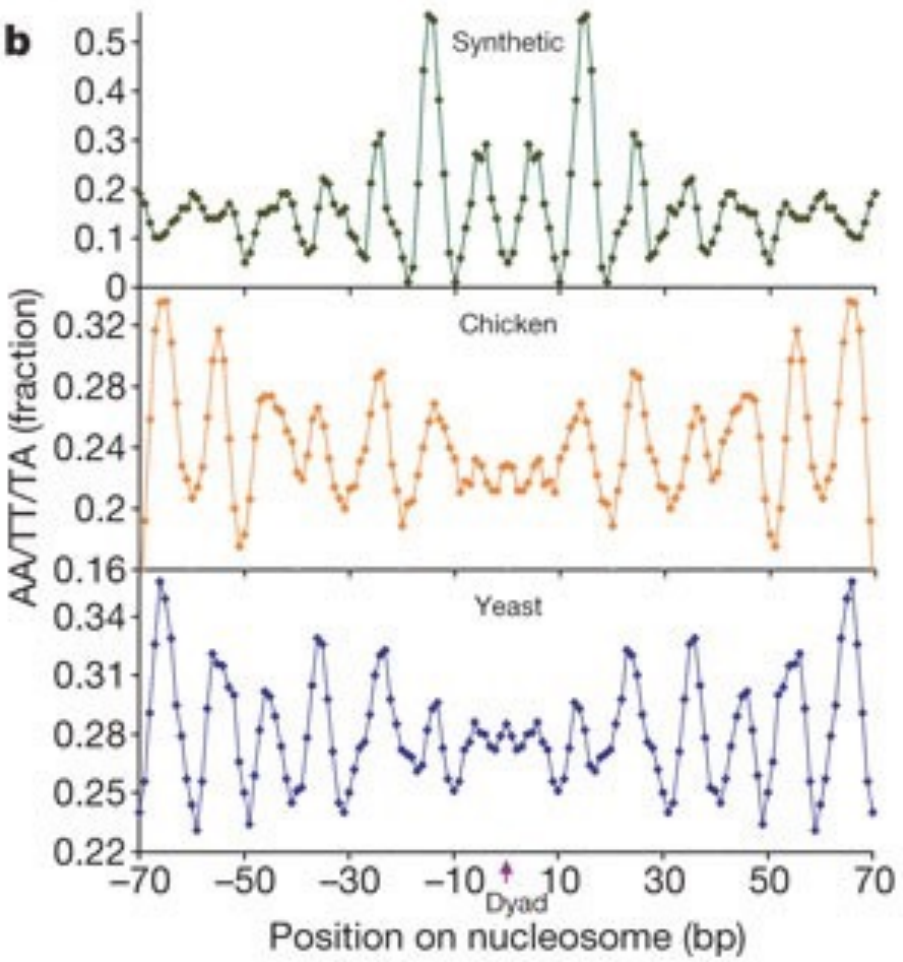
- Nucleosomes can inhibit *in vitro* transcription initiation (Knezetic and Luse 1986)
- However, actively transcribing RNA PolII can displace nucleosomes (LaPointe and Kornberg 1987)



Knezetic and Luse (1986)

Chromatin Structural Specificity

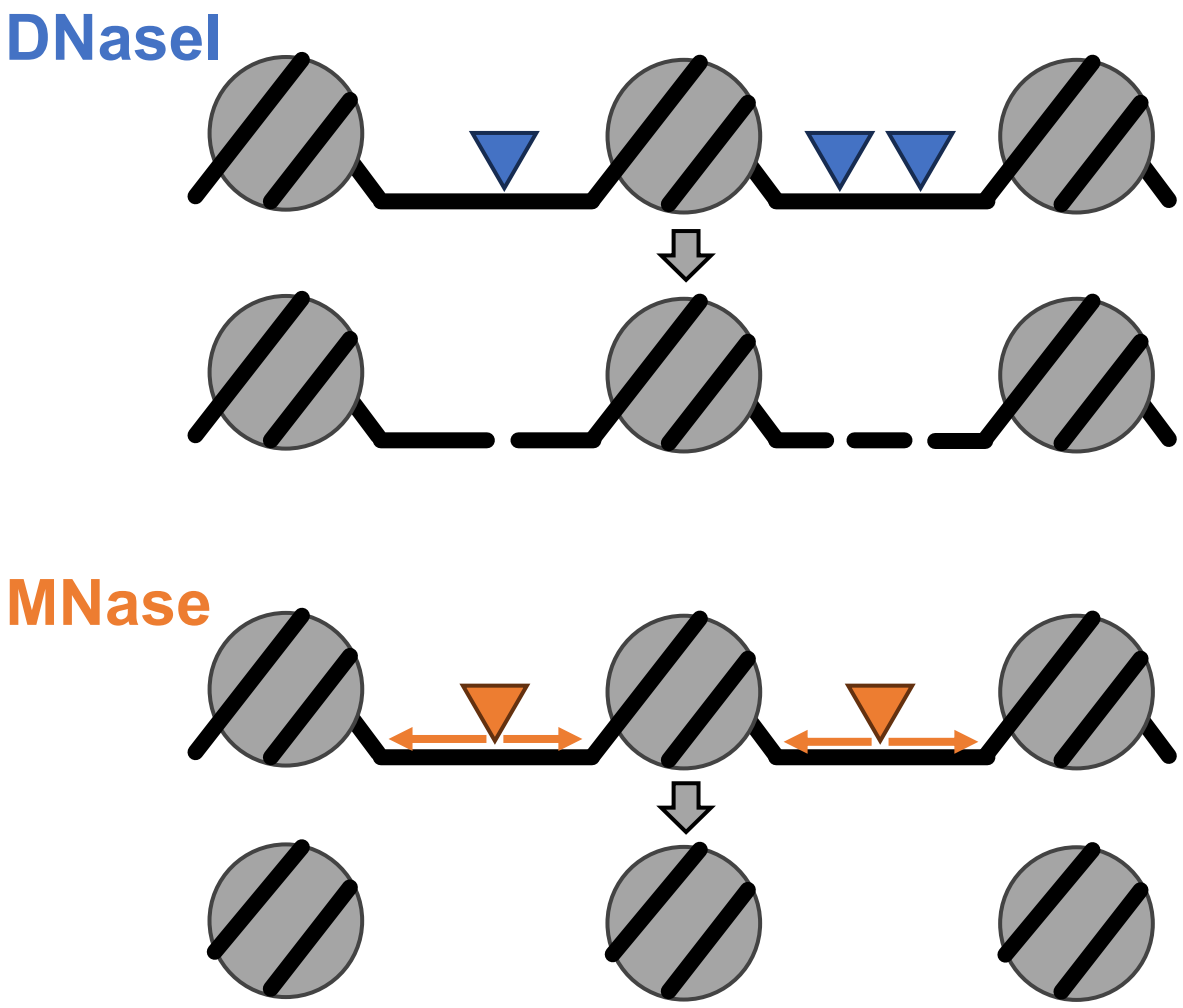
- Major question: **Can chromatin accomplish sequence-specific inhibition of transcriptional activity?**
- Nucleosomes do not have protein-oriented sequence specificity
- BUT DNA sequences can be more or less energentically favorable for nucleosomal wrapping (Segal…Widom 2006)



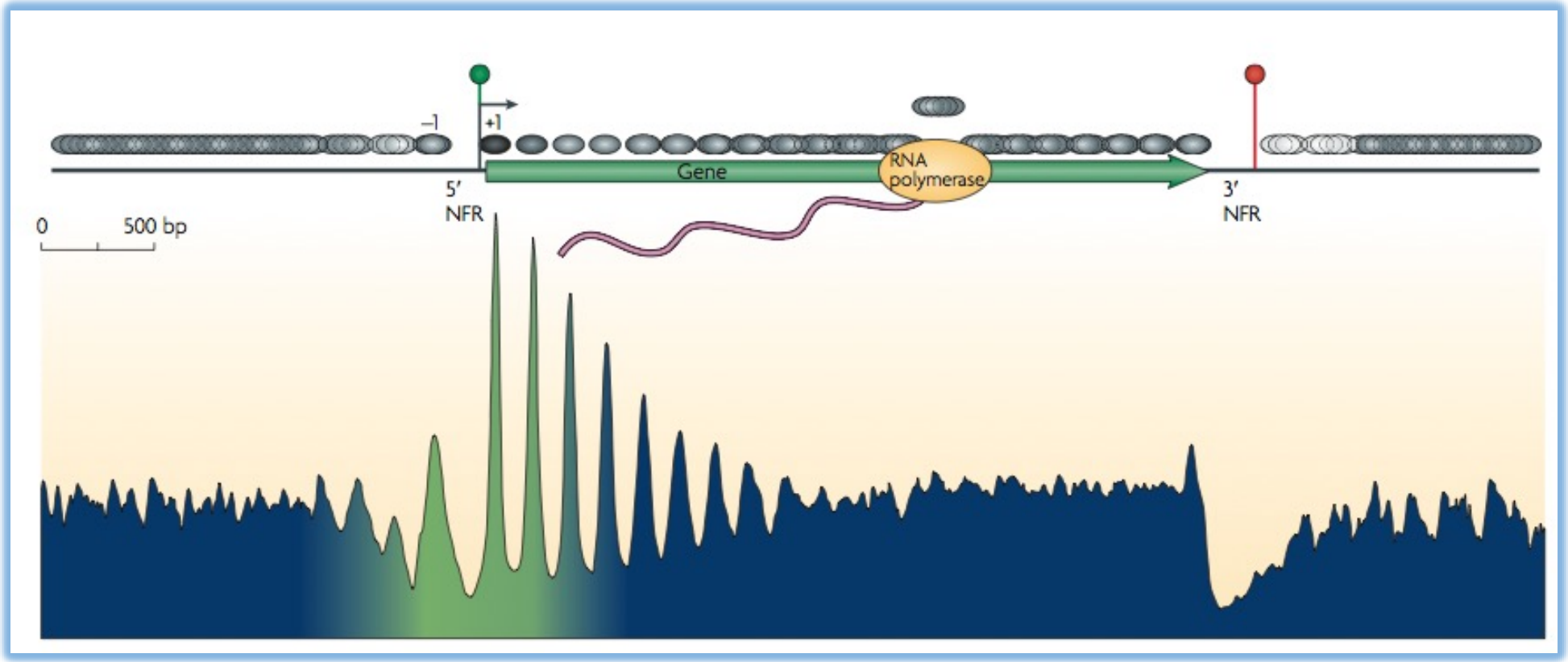
Segal et al. (2006)

Mapping nucleosome positions

- DNaseI
  - Endonuclease
  - Digests in nucleosome depleted regions
  - DNA ends represent **lack of nucleosomes**
- Micrococcal Nuclease (MNase)
  - Processive endo/exonuclease
  - Nucleosomal DNA is protected
  - Residual DNA represents **presence of nucleosomes**



Genomic distribution of nucleosomes

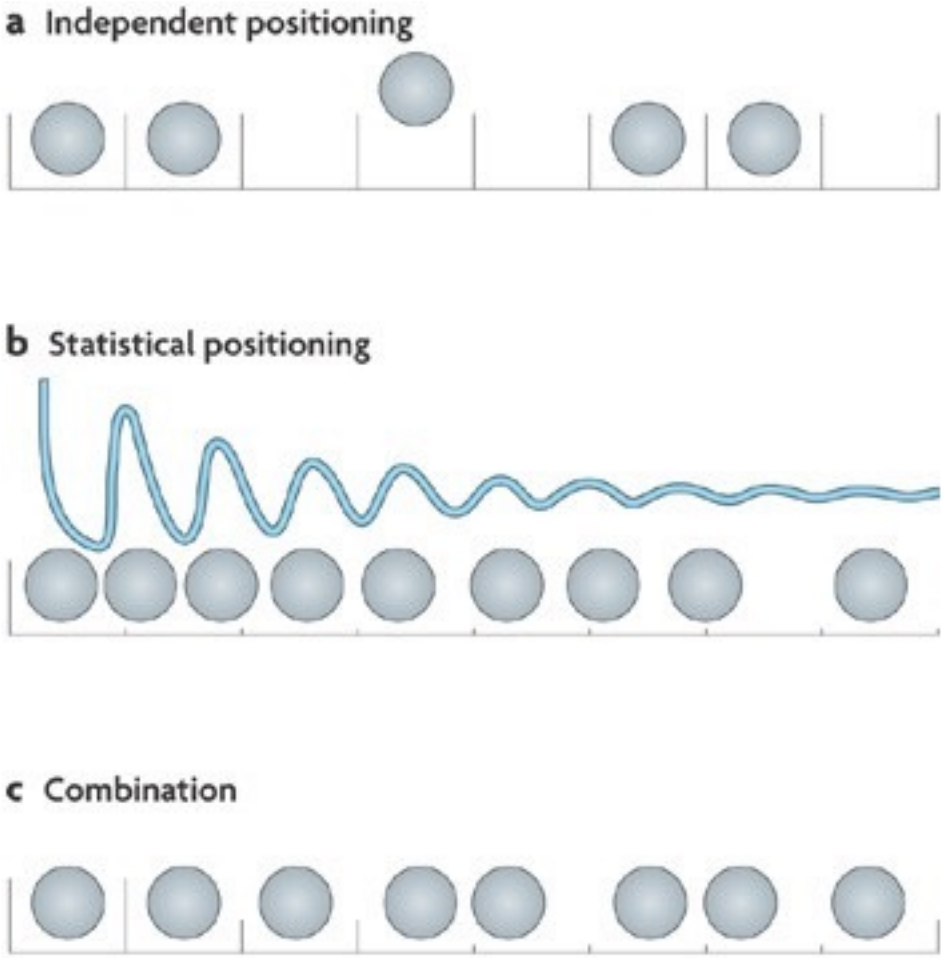


- The presence of NFRs demonstrated that open promoter states are stable and common, even at genes that are transcribed so infrequently

Cizhong Jiang, B Franklin Pugh, Nature Review Genetics, 2009 vol. 10 (3) pp. 161-72

Nucleosome phasing occurs around transcription start sites

- Which comes first, nucleosome positioning or RNA polymerase?
- Model 1: Nucleosomes are **static**, proper positioning is **causal** for RNA polymerase loading: “Independent positioning”
- Model 2: Nucleosomes are **dynamic**, positioning is a **consequence** of RNA polymerase loading: “Statistical positioning”

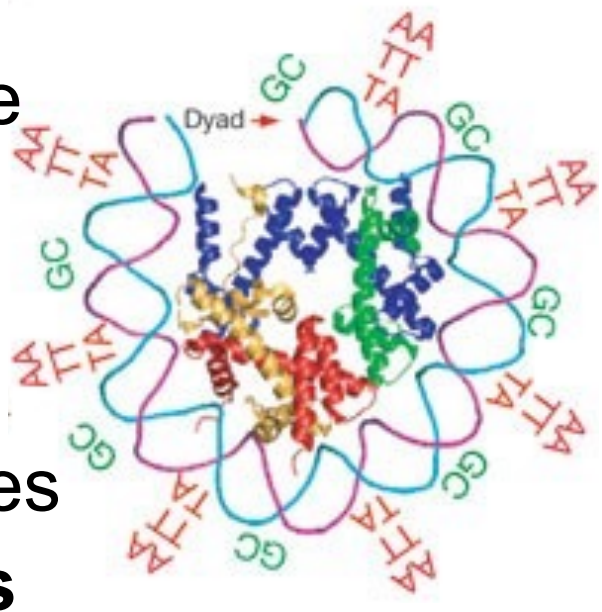


Jiang and Pugh (Nat. Rev. Genet. 2009)

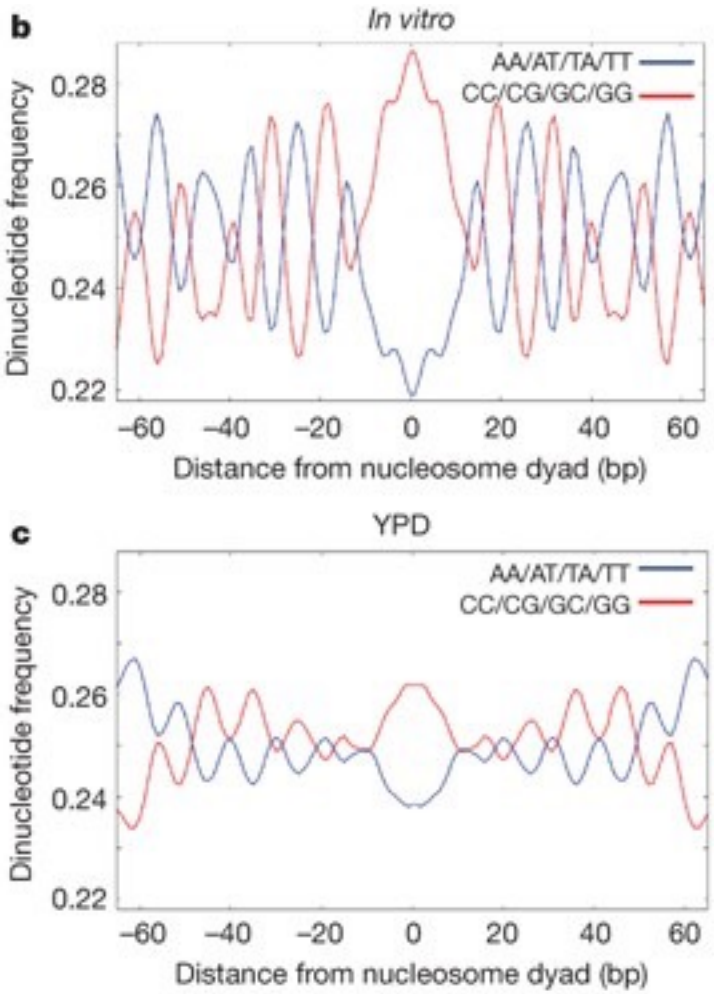
Nature Reviews | Genetics

Chromatin Structural Specificity

- Comparing nucleosomal affinity *in vitro* vs. *in vivo* in yeast shows lower sequence preference *in vivo* (Kaplan…Segal 2009)
- Inherent sequence preferences appear even lower in mammalian genomes
- **Nucleosome positioning is not inherently heritable by sequence specificity**



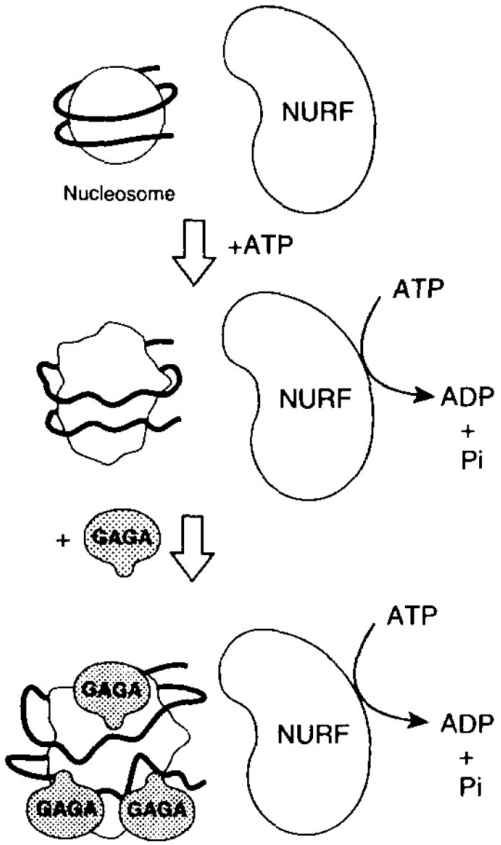
Segal et al. (2006)



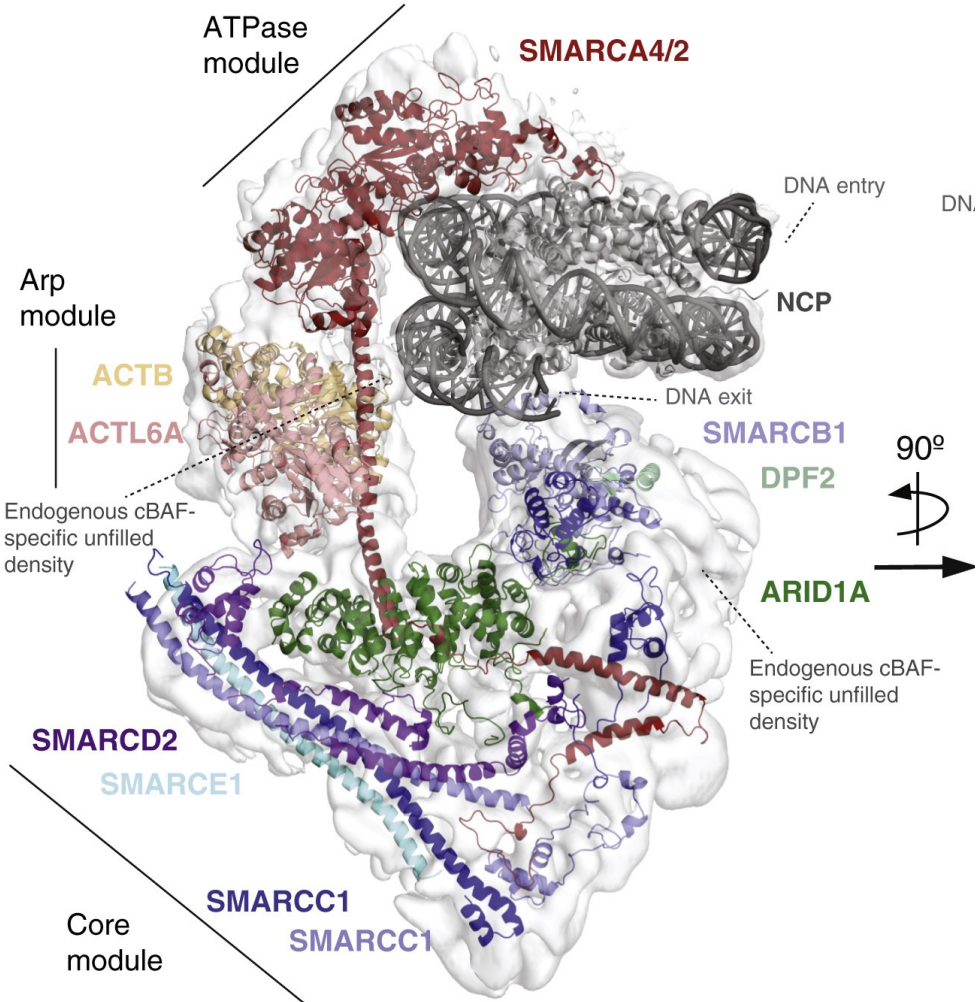
Kaplan et al. (2009)

Chromatin remodelers can shift and evict nucleosomes from the DNA template

- SWI/SNF: Sliding and eviction (BAF in mammals)
- ISWI: Nucleosome spacing
- Remodelers can play different roles in mammals vs. yeast—yeast have strong sequence-defined nucleosome-depleted regions (NDRs)
- Nucleosome positioning = Sequence preference + remodeler activity + protein barriers (e.g. RNA polymerase)



Tsukiyama…Wu (1996)

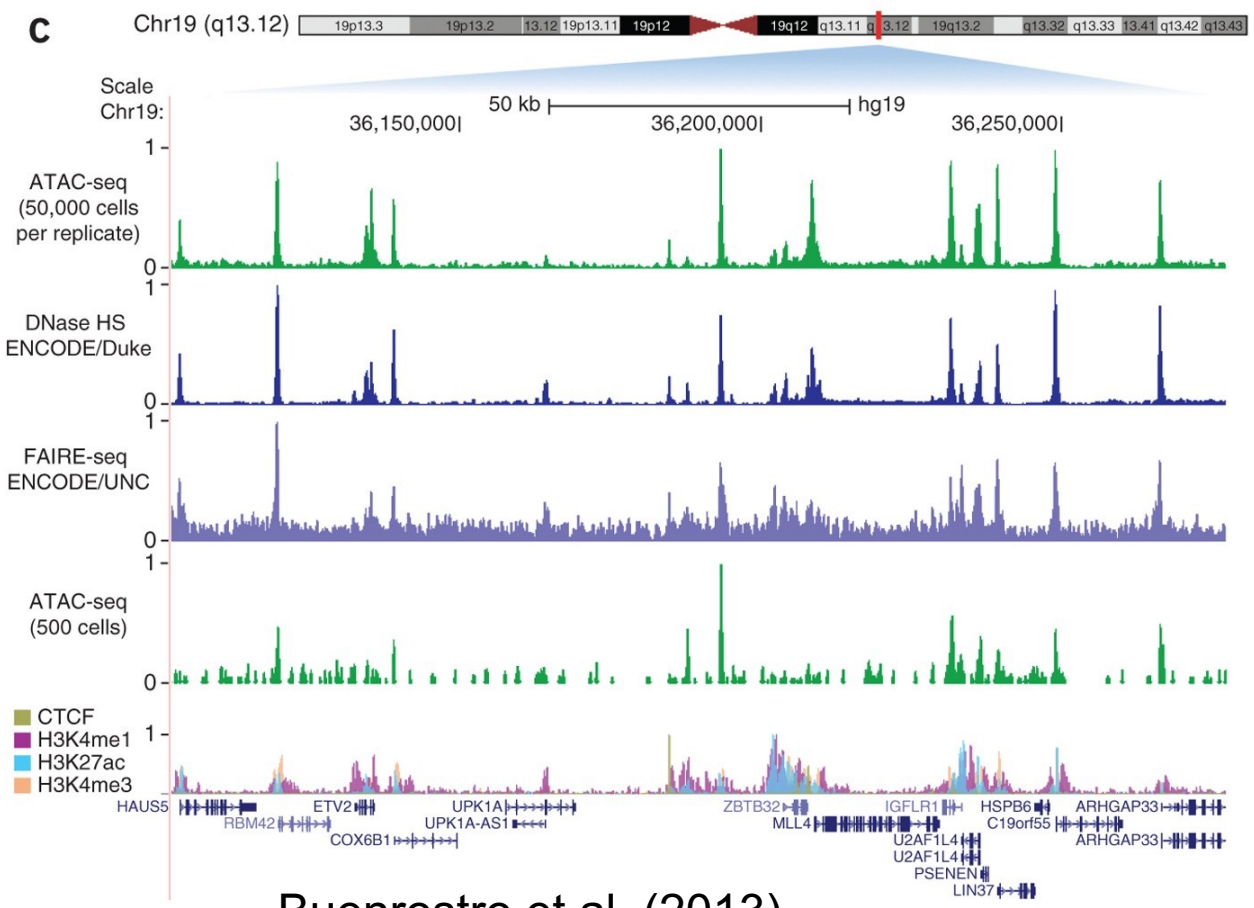


Mashtalir…Kadoch (2020)

Is nucleosome positioning **epigenetic**?

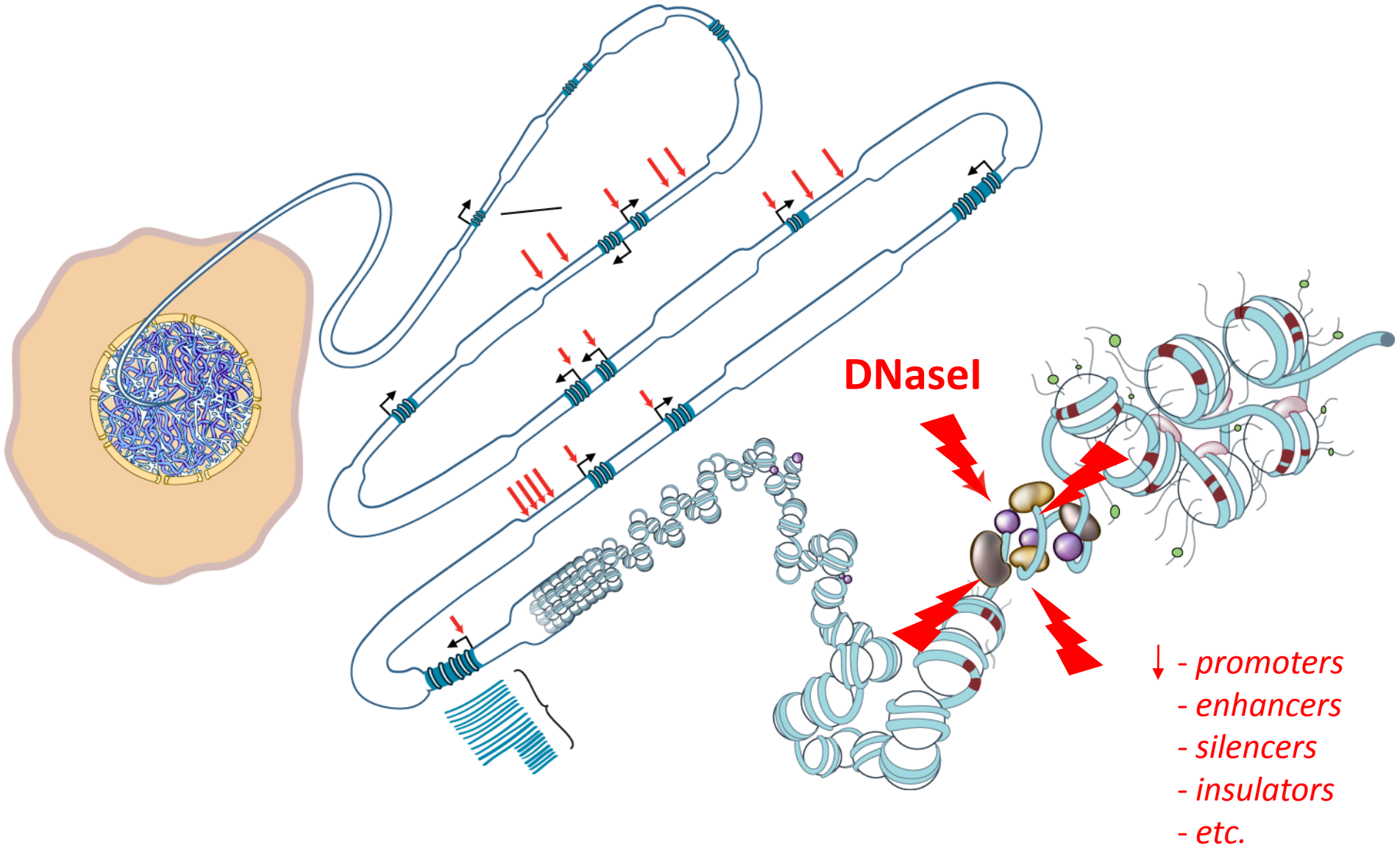
Chromatin Accessibility

- **Chromatin accessibility** is inherently associated with transcriptional activity
- Chromatin accessibility profiling by high-throughput sequencing :
  - **MNase-seq** (nucleosome occupancy) (Fire 2006)
  - **FAIRE-seq** (Lieb 2007)
  - **DNase-seq** (Crawford 2008)
  - **ATAC-seq** (Greenleaf 2013)



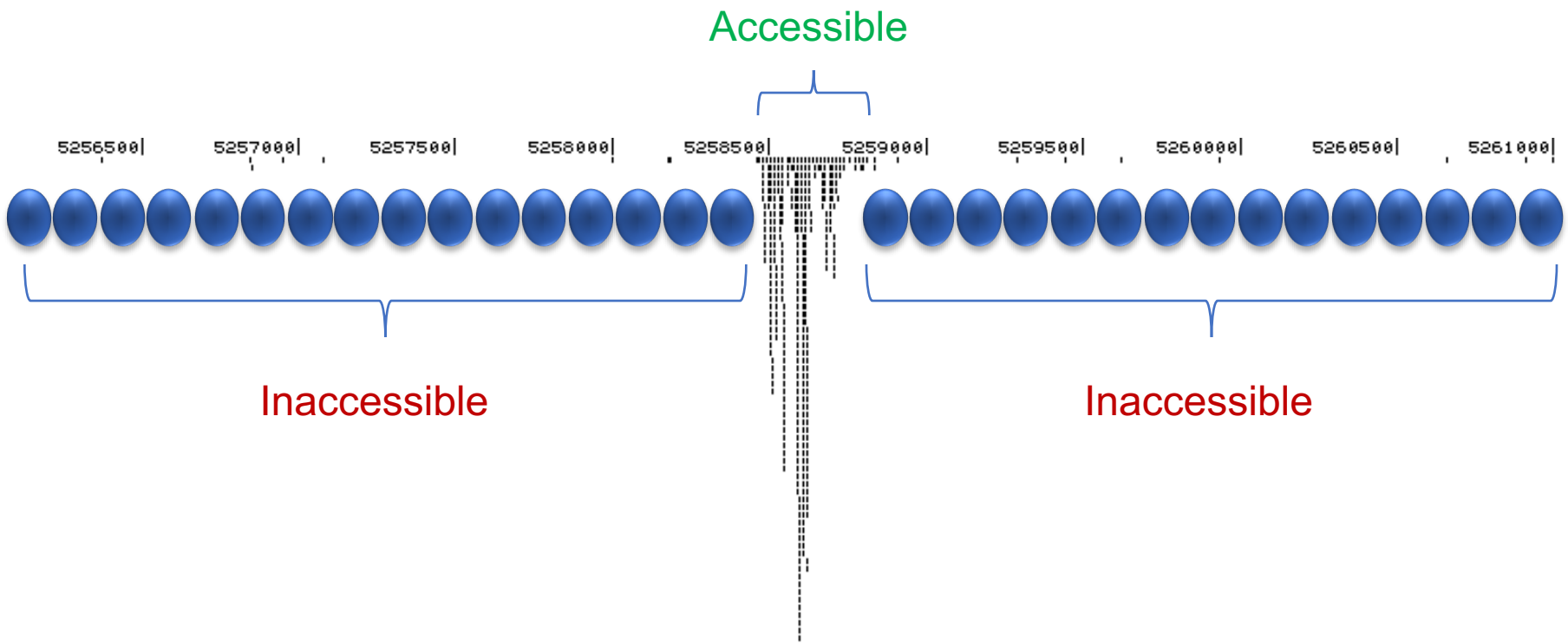
Buenrostro et al. (2013)

DNase I hypersensitivity ~ Regulatory DNA

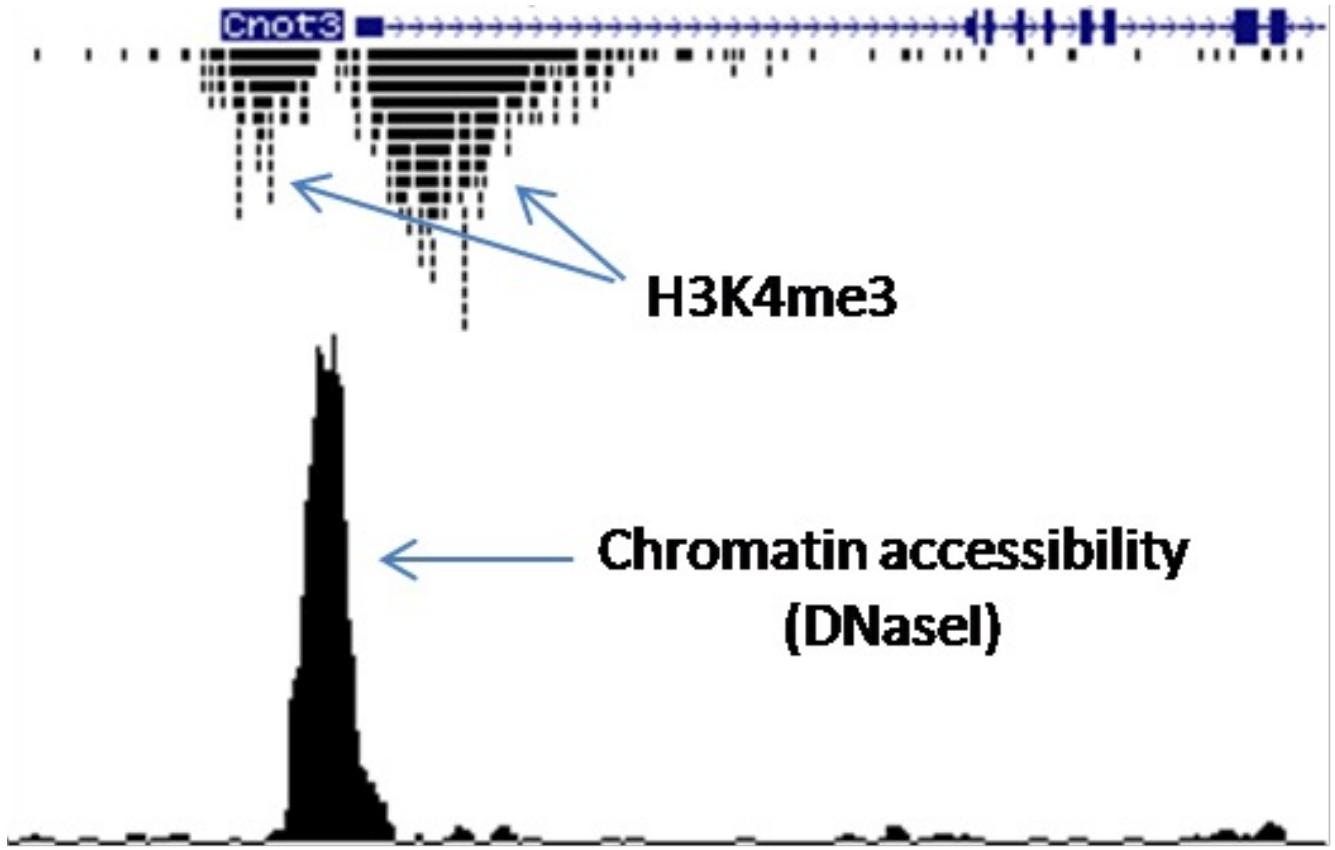


Digital DNaseI profiling

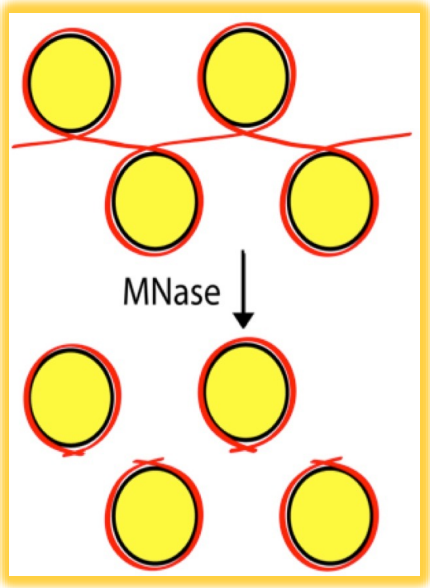
Precise delineation of the accessible regulatory DNA compartment

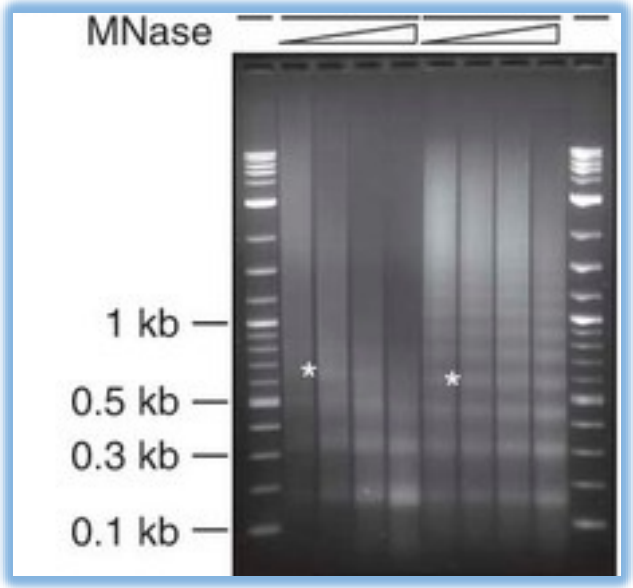


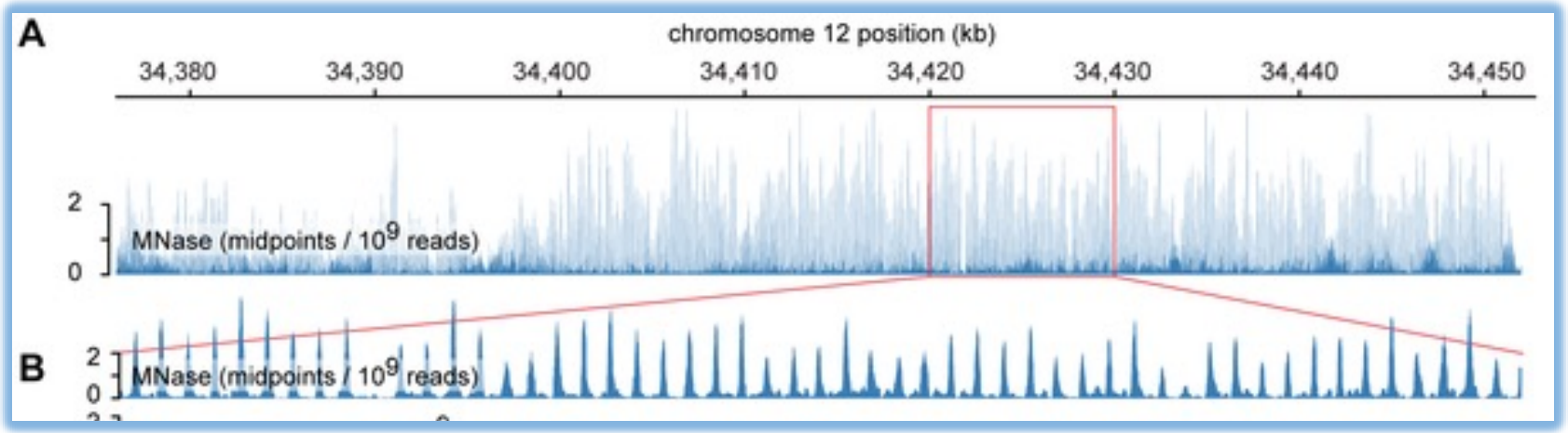
Digital DNaseI profiling: direct access to regulatory sequences



MNase-seq

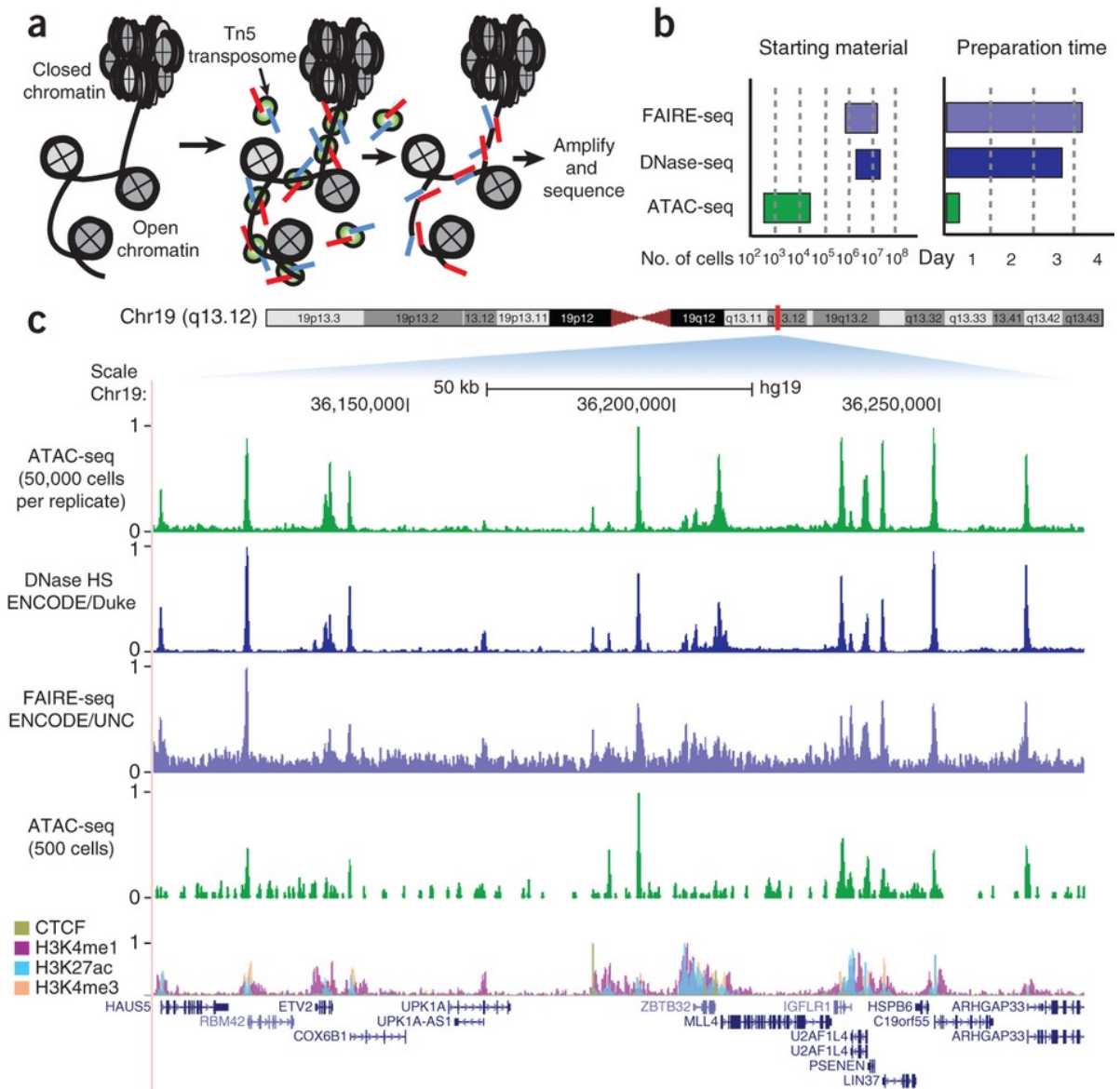




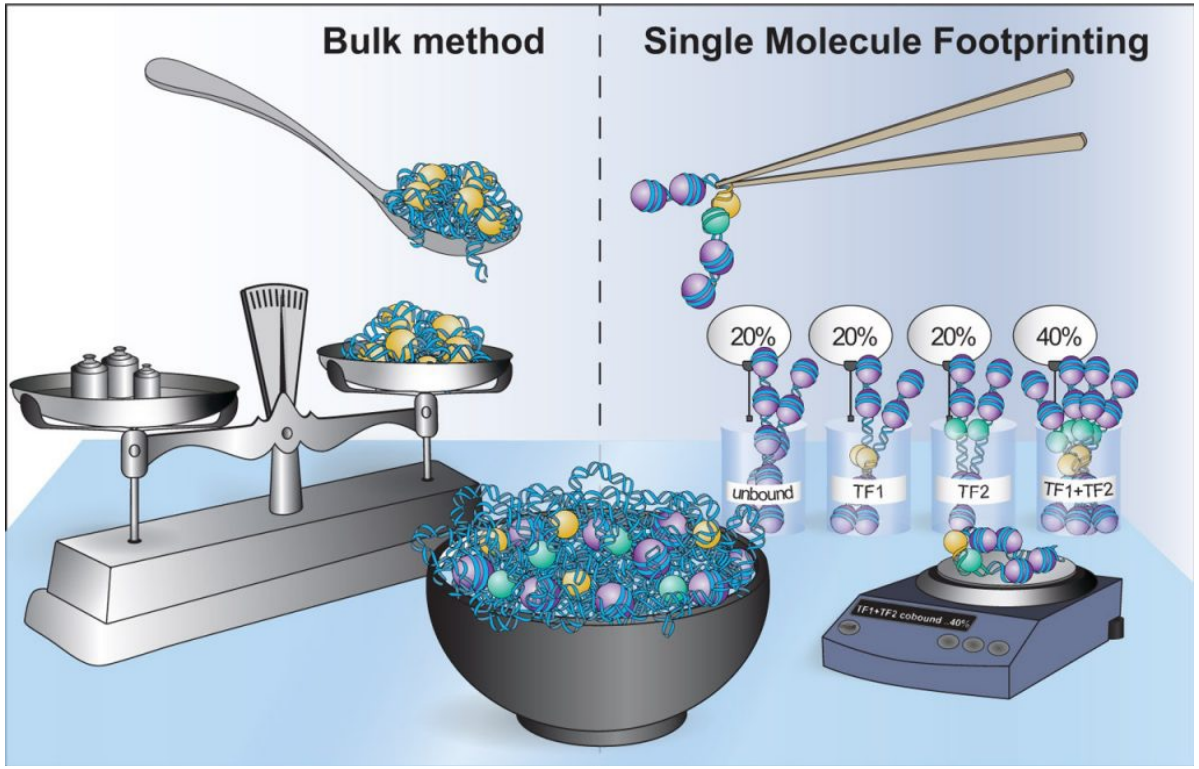


Gaffney DJ, McVicker G, Pai AA, Fondufe-Mittendorf YN, et al. (2012) Controls of Nucleosome Positioning in the Human Genome. PLoS Genet 8(11): e1003036. doi:10.1371/journal.pgen.1003036
http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1003036

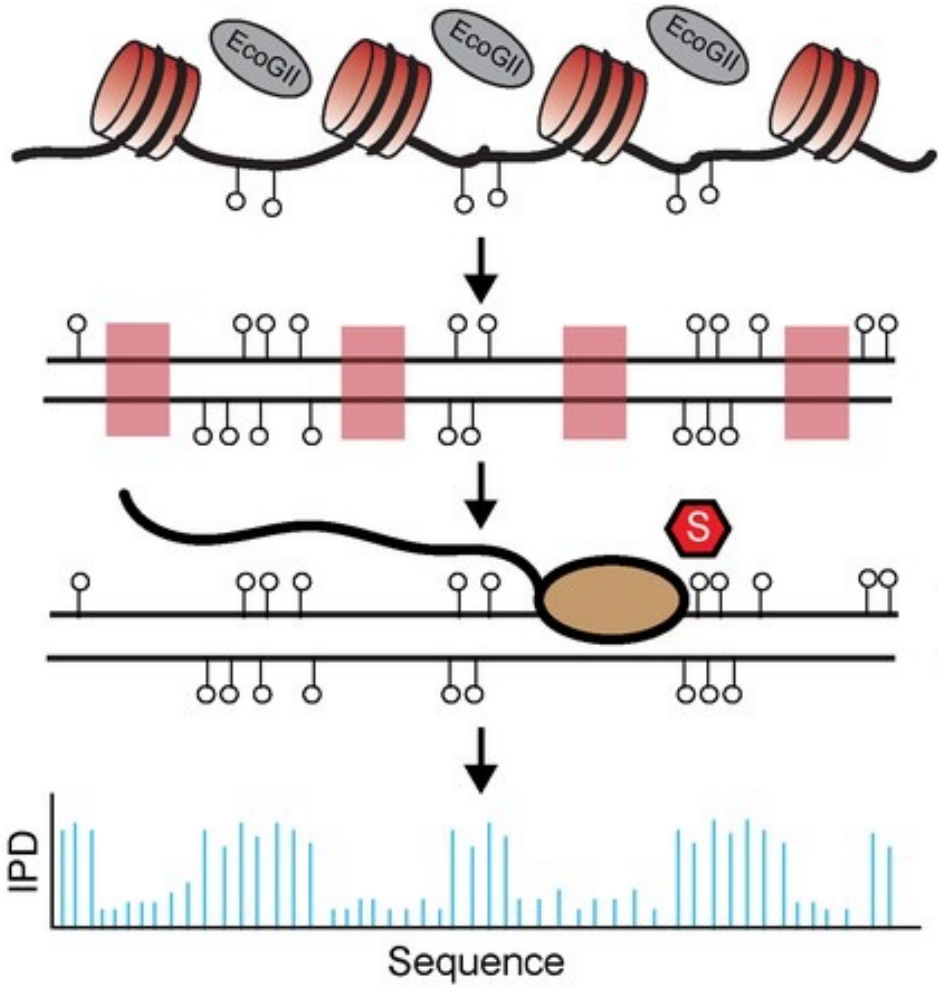
ATAC-seq
(assay for transposase-accessible chromatin)



Single molecule footprinting: Detecting multiple nucleosomes on the same DNA molecule



Arnaud Krebs Lab



Abdulhay…Ramani (2020)

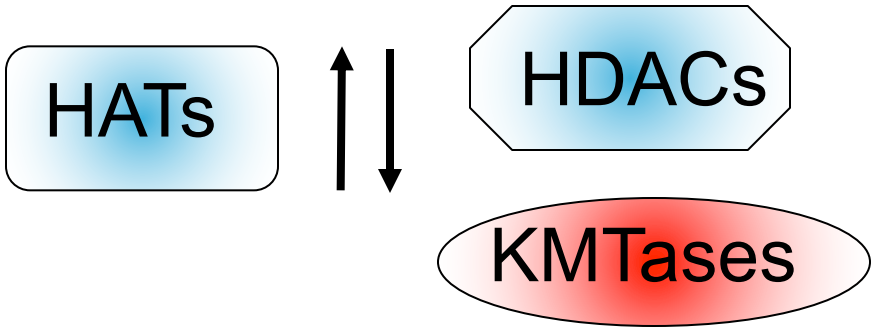
Post-translational Histone Modifications

H3 tail Modifications:

=Acetylation

=Methylation

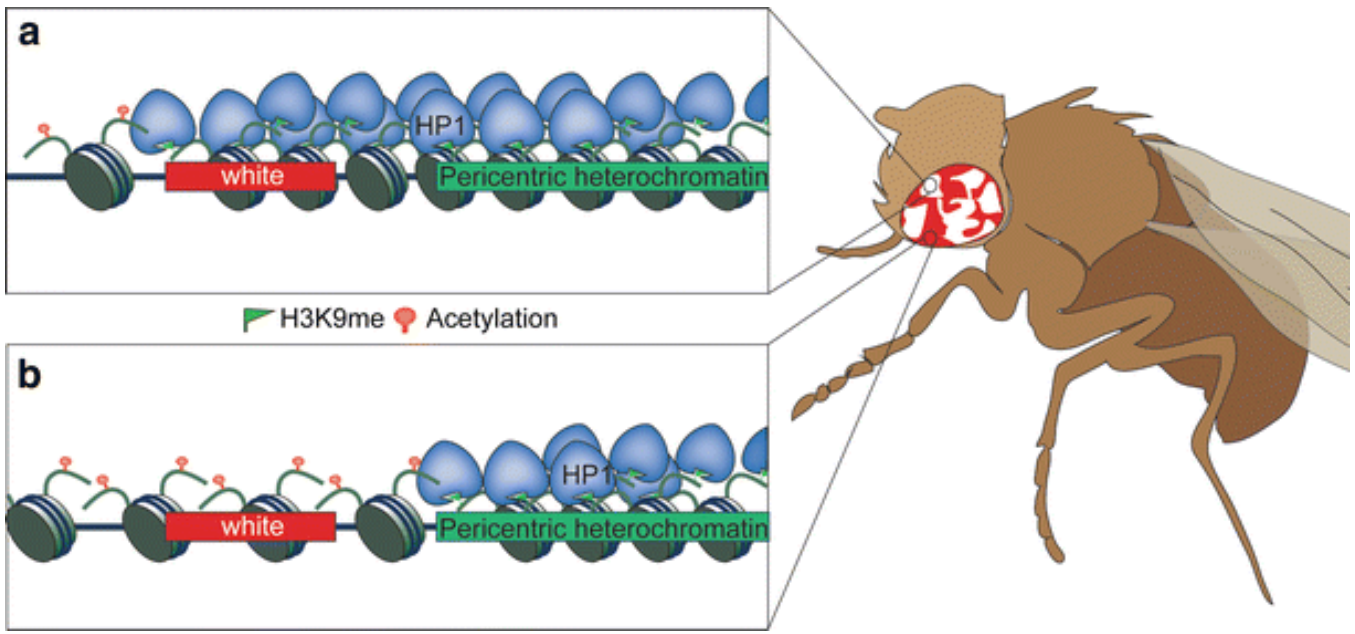
Active



Repressive

History of histone modifications

- Discovery of modifiers of position-effect variegation in *Drosophila* eyes (e.g. Su(var)3-9)
- Discovery of histone acetyltransferase in tetrahymena with homology to yeast Gcn5 (Brownell…Allis 1996)
- Discovery that homologs of Su(var)3-9 are histone methyltransferases (Late 90s)
  - That methylate H3K9!



Wang et al. (2014)

Chromatin Structure: More History

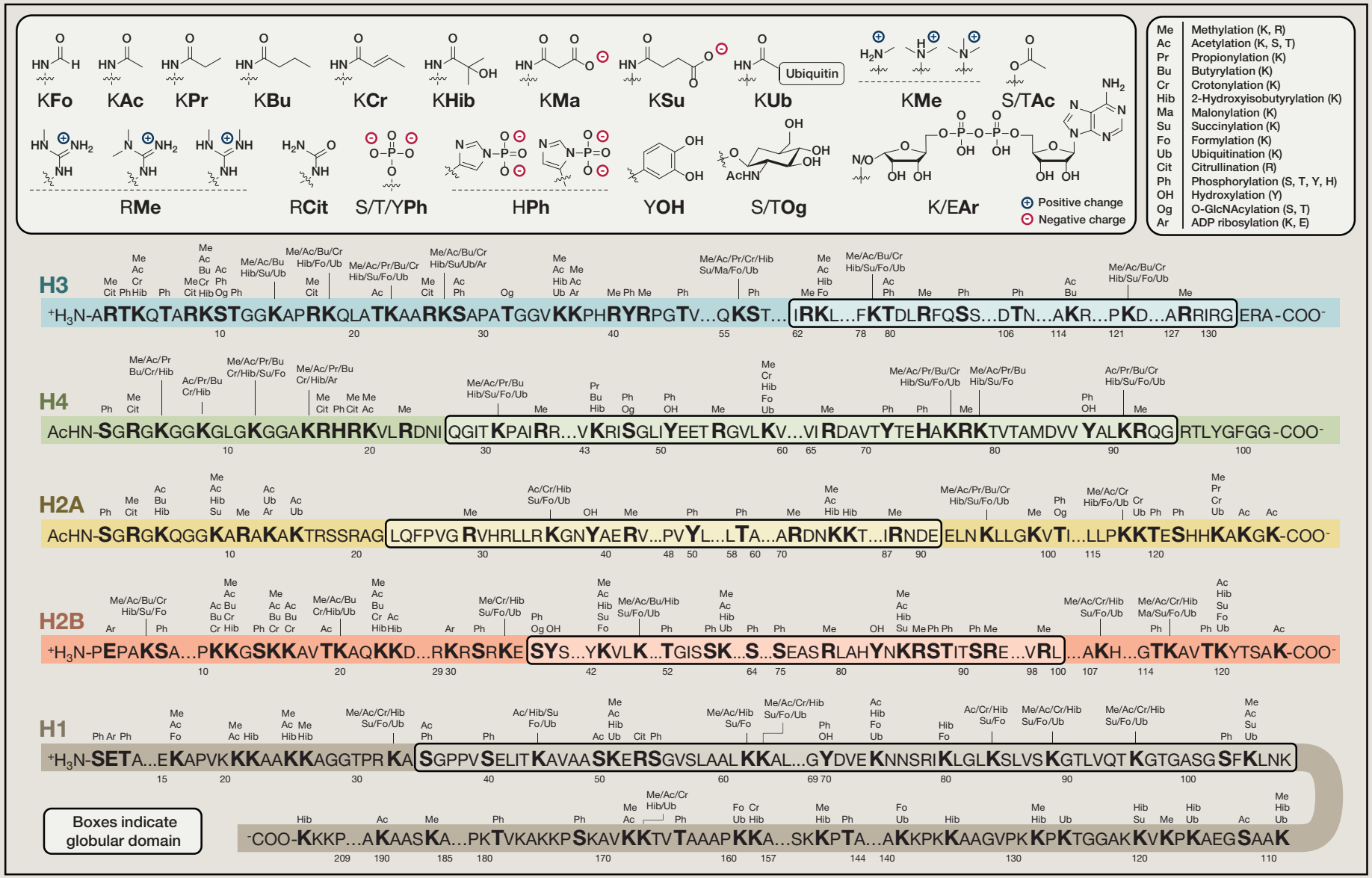
- Condensed chromatin can adopt compact inaccessible fiber structure (Woodcock 1998)
- Chromatin condensation is associated with specific histone post-translational modifications such as H3K9me3 (Jenuwein 2000) and H3K27me3 (Zhang 2002)
- The latter is associated with Polycomb repressive complexes that physically condense chromatin (Woodcock 2004)

Bednar et al. (1998)

458 Cell 159, October 9, 2014 ©2014 Elsevier Inc. DOI http://dx.doi.org/10.1016/j.cell.2014.09.037

SnapShot: Histone Modifications

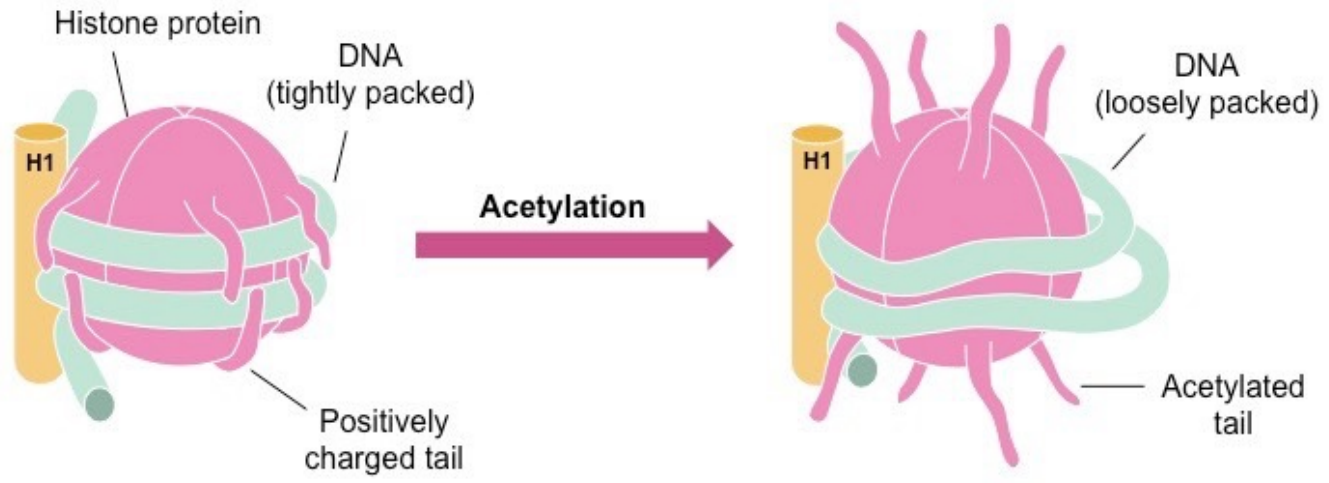
He Huang,[1] Benjamin R. Sabari,[2] Benjamin A. Garcia,[3] C. David Allis,[2] and Yingming Zhao[1]
[1]Ben May Department of Cancer Research, The University of Chicago, Chicago, IL 60637, USA
[2]Laboratory of Chromatin Biology and Epigenetics, The Rockefeller University, New York, NY 10021, USA
[3]Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104, USA



See online version for for legend and references.

Histone acetylation: Charge neutralization “loosens” DNA

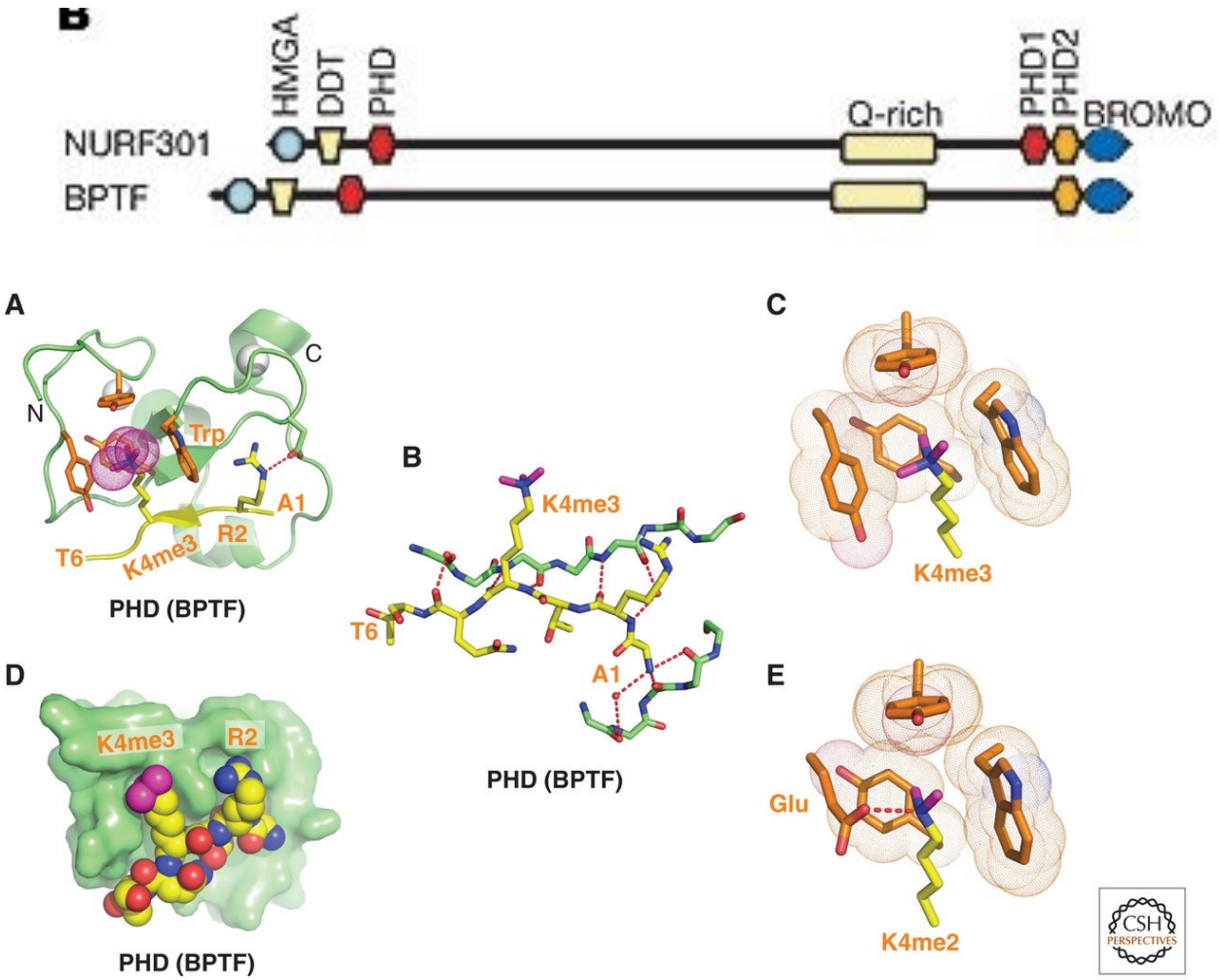
- Understood as early as 1964 (Vincent Allfrey)
- Many acetylated lysines in histones are enriched at active promoters/ enhancers
- However, most histone PTMs do not change the charge of the modified amino acid—how do they do what they do?



Bioninja

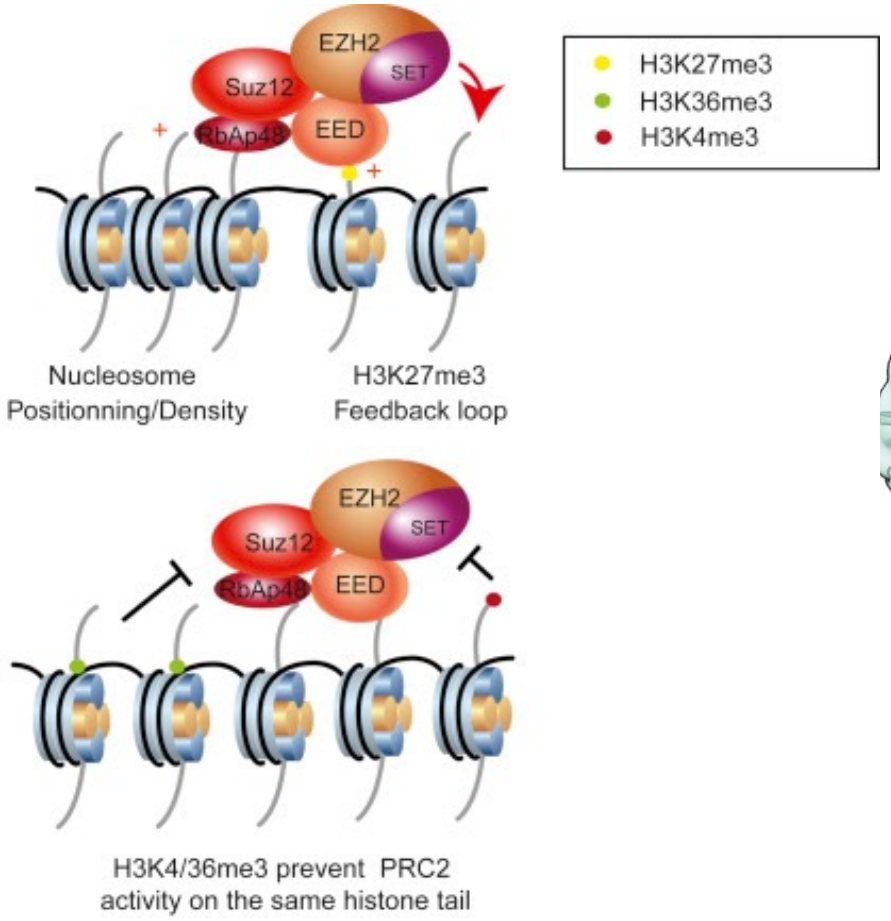
Histone modifications can act as binding platforms

- Specific reader domains can interact with specific PTMs
  - Bromodomain = Acetyl-lysine
  - Chromodomain = Methyl-lysine
  - PHD finger = H3K4me
  - Royal domain = Methyl-lysine
- Example: BPTF PHD + Bromodomain engages both H3K4me and H3Kac

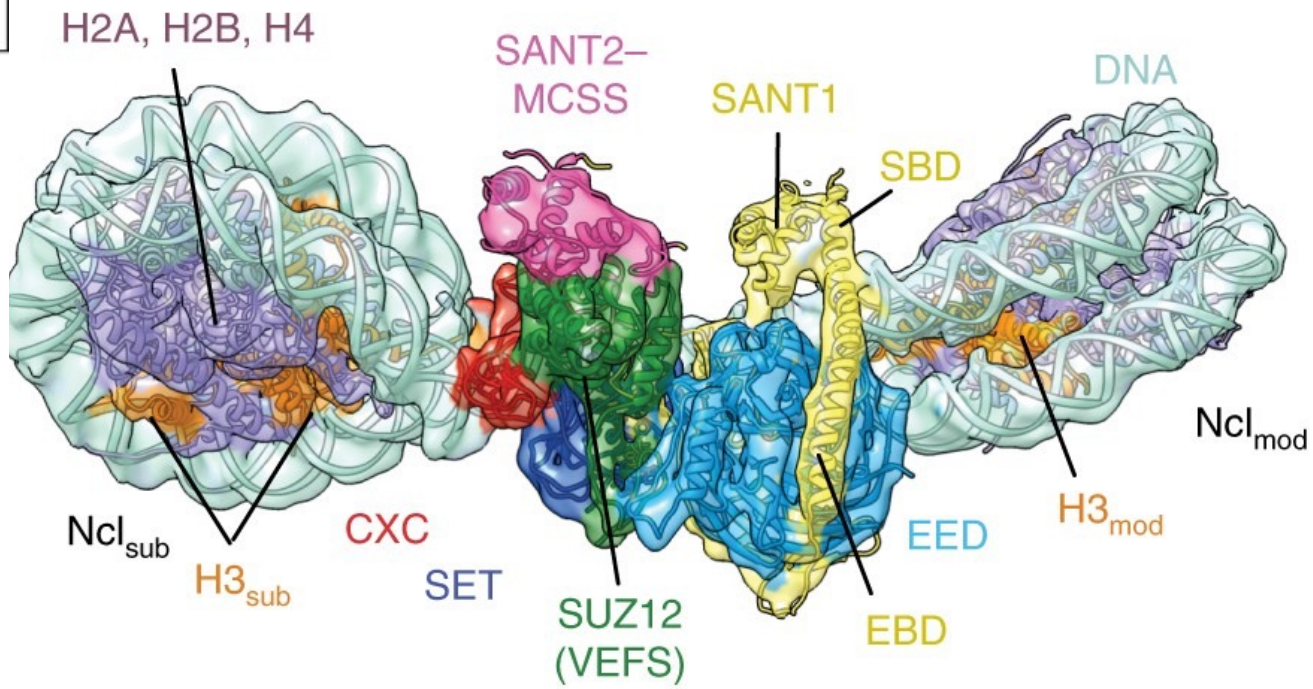


Dinshaw Patel

Heterochromatin read-write machanisms



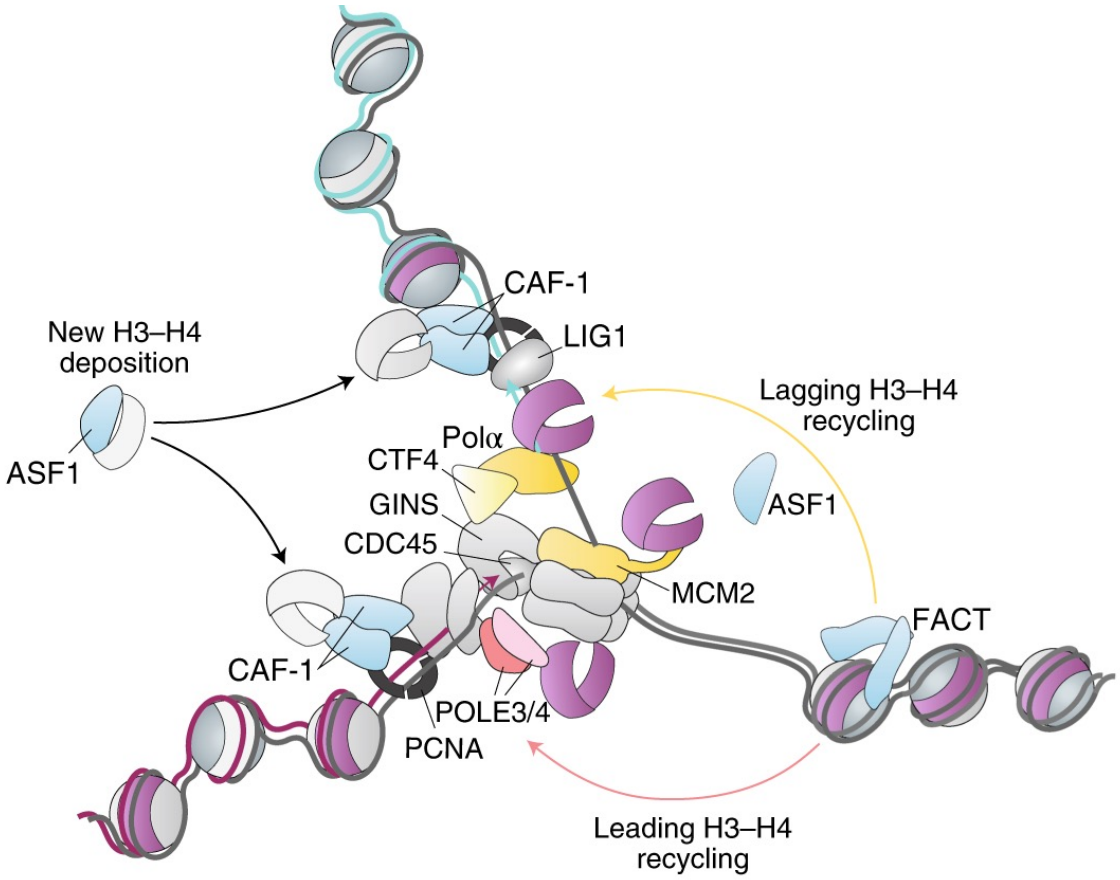
Holoch and Margueron (2017)



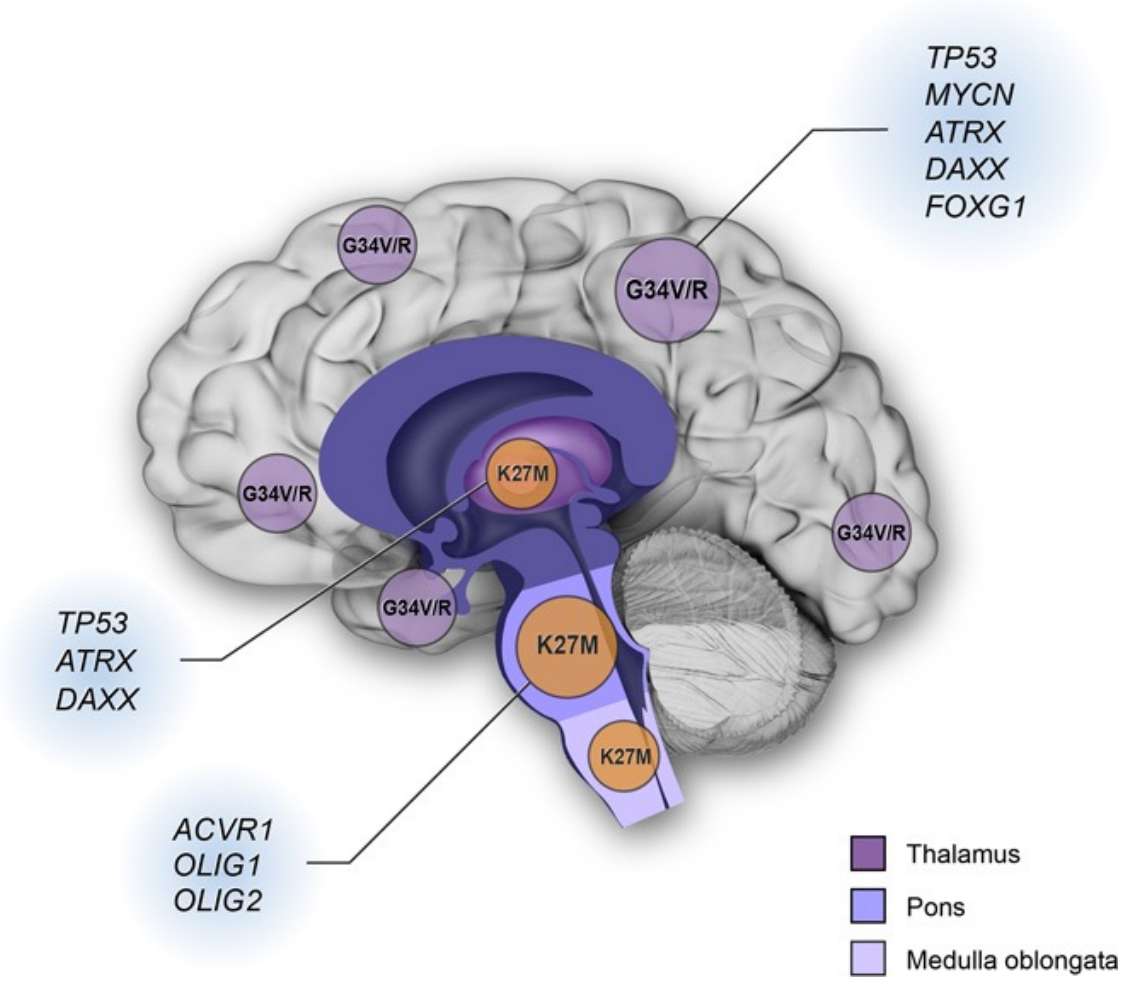
Poepsel…Nogales (2018)

Replication disrupts chromatin during each cell cycle

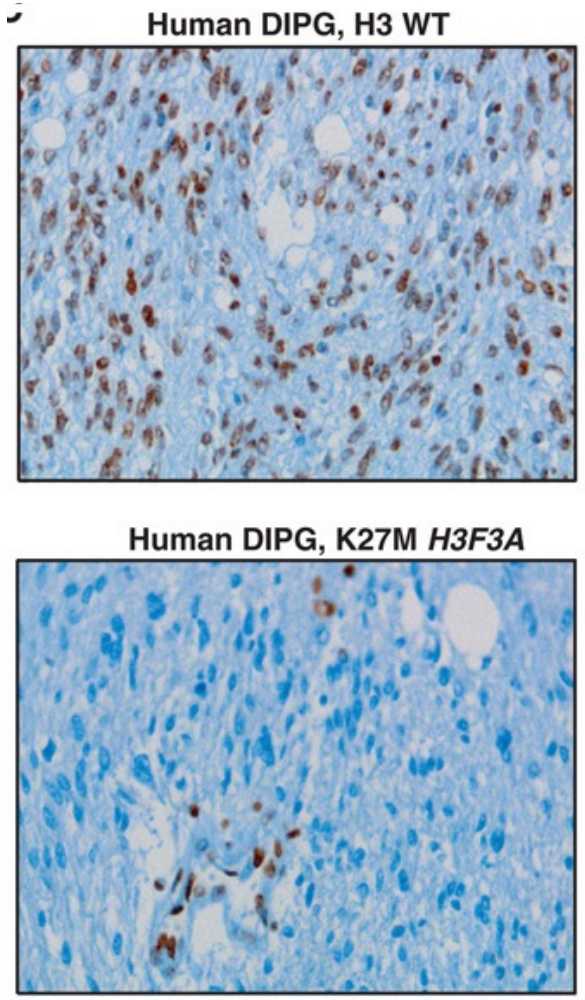
- Nucleosomes are partitioned equally to daughter strands by MCM2 histone chaperone activity (Petryk…Groth 2018)
- Heterochromatic histone modifications are well positioned after replication, but euchromatic modifications are not! (Escobar…Reinberg 2019)



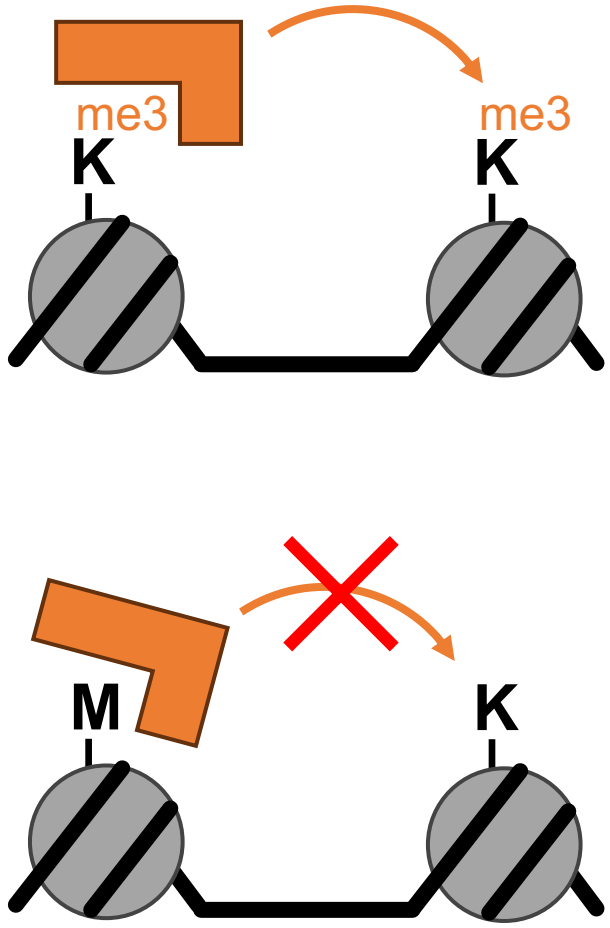
Mutations in core histones cause rare pediatric cancers



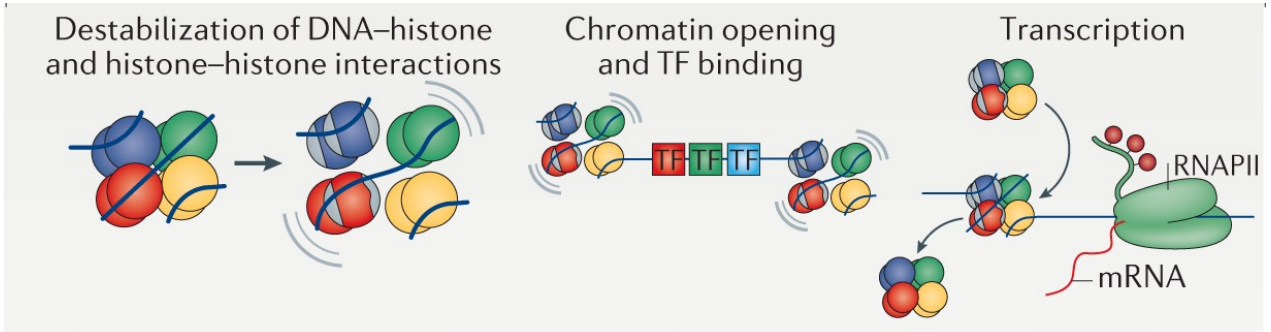
Lulla et al. (2016)



Lewis…Allis (2013)



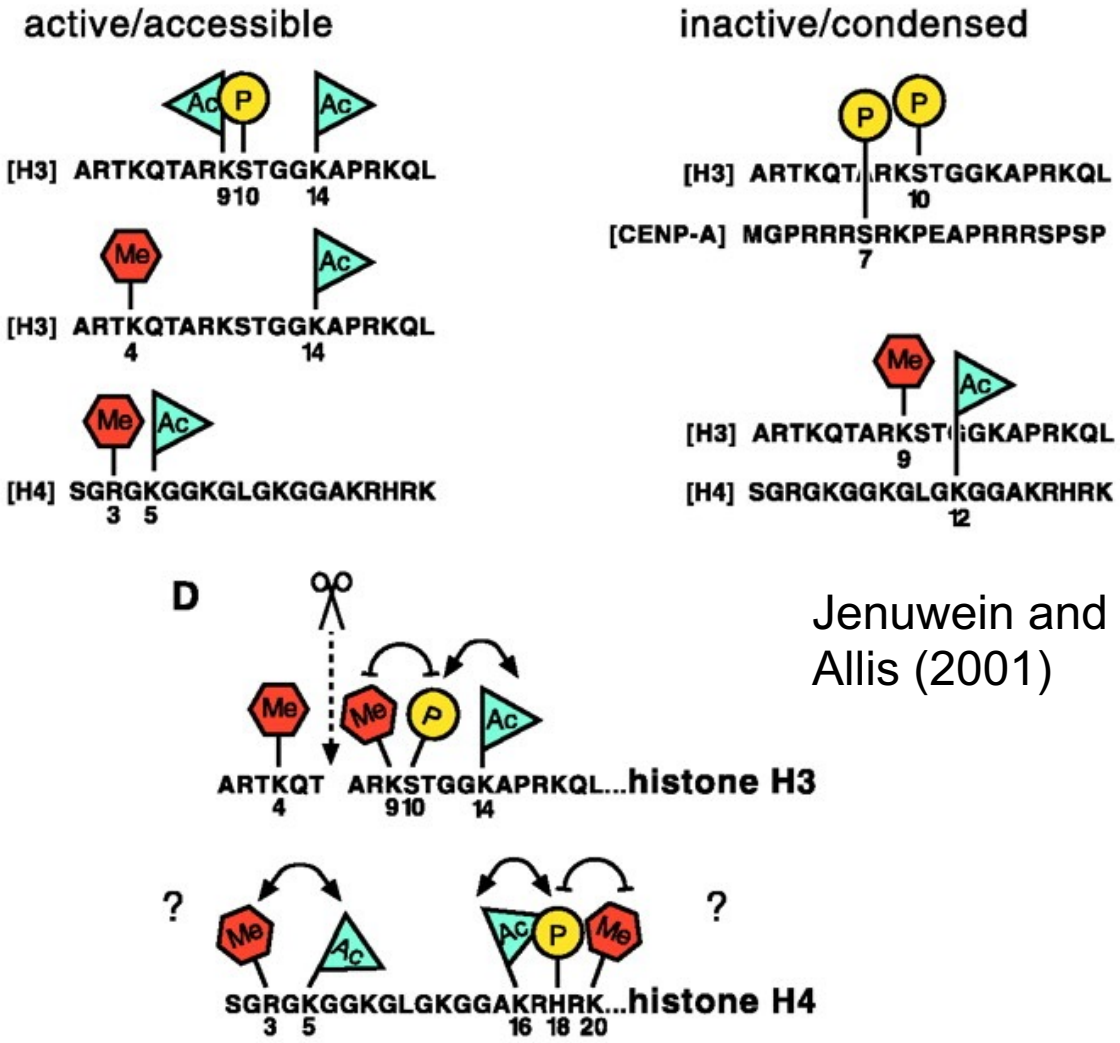
Histone variants modify nucleosome function



- H2A
  - H2A.X = at DNA damage sites
  - H2A.Z = Active tx. regions
  - MacroH2A = compacted heterochromatin
- H3
  - H3.3 = Deposited at active tx. sites AND telomeres (different histone chaperones)
  - CENP-A (CenH3) = Centromeres
  - Testis-specific H3 variants = Nucleosome destabilization

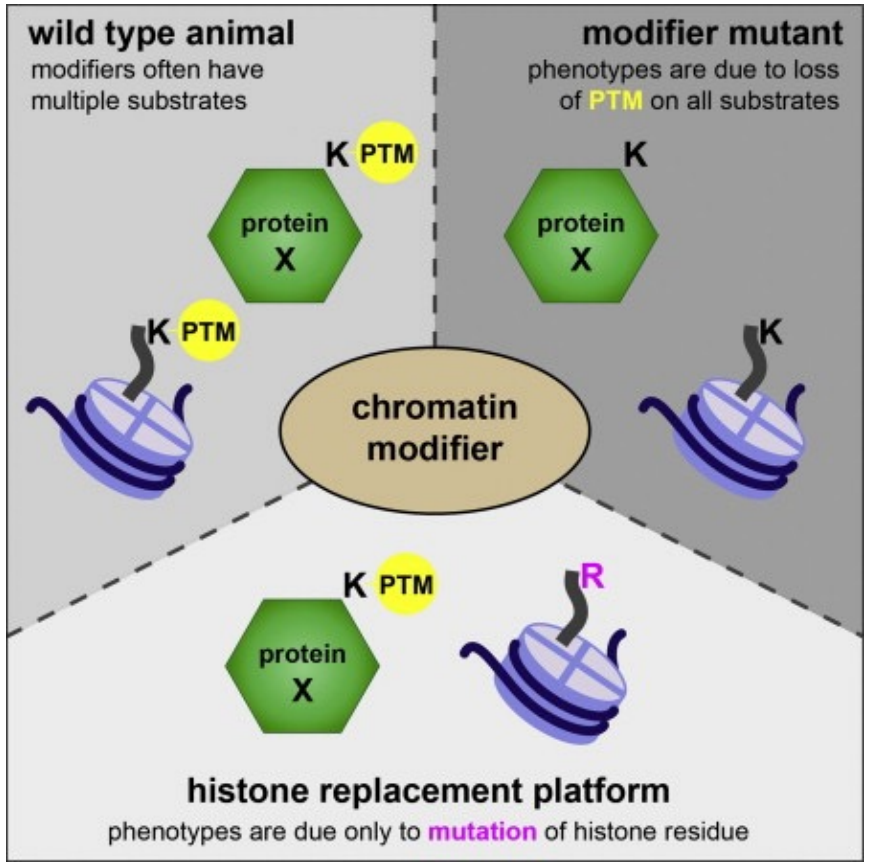
Are chromatin accessibility/modifications "epigenetic"?

- The "histone code" hypothesis
- "*We propose that this epigenetic marking system represents a* ***fundamental regulatory mechanism*** *that has an impact on most, if not all, chromatin-templated processes, with far-reaching* ***consequences for cell fate decisions and both normal and pathological development***"

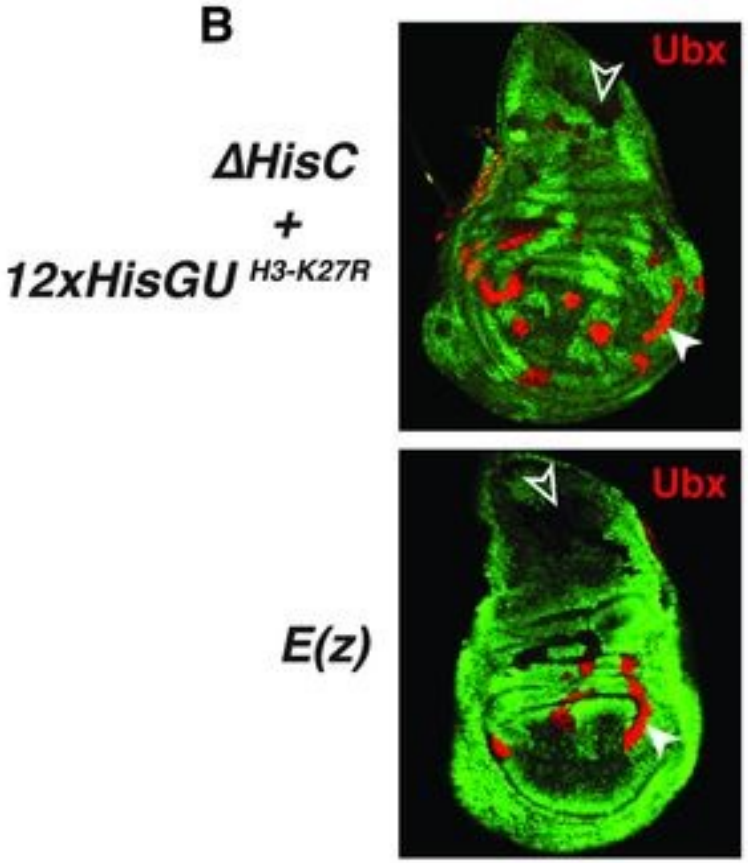


Jenuwein and Allis (2001)

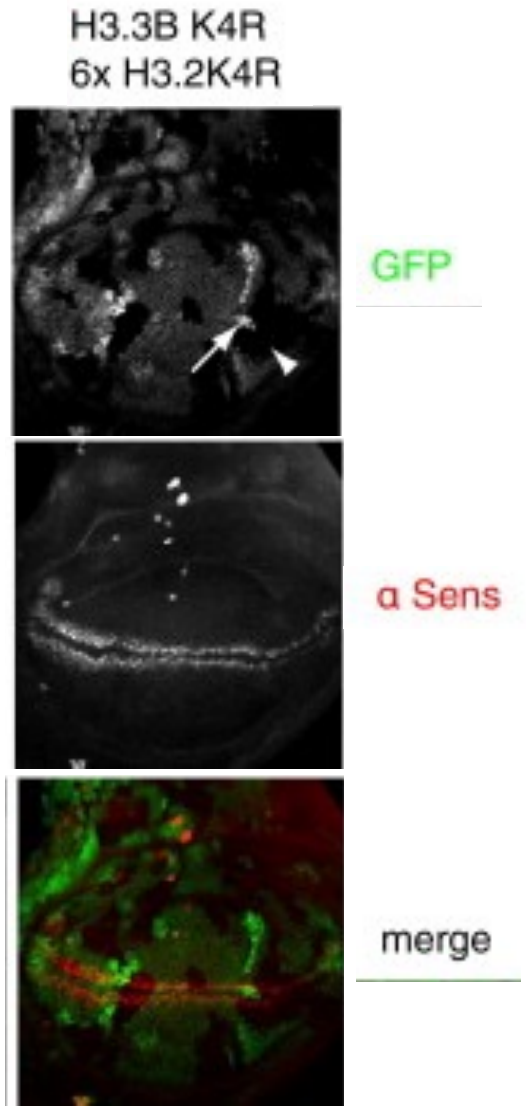
Direct histone mutations in *D. melanogaster* suggest some histone modifications are dispensable for transcription/viability



McKay et al. (2015)

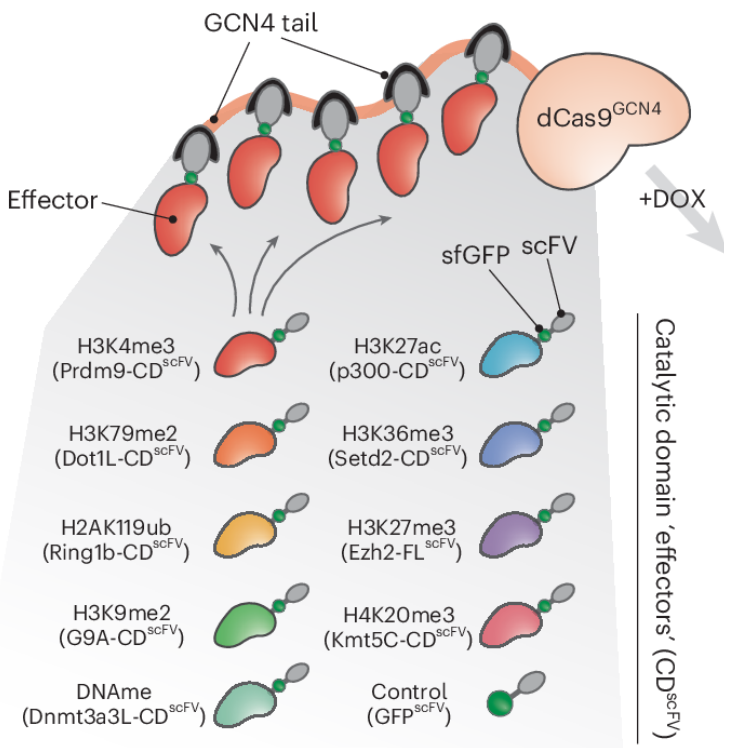


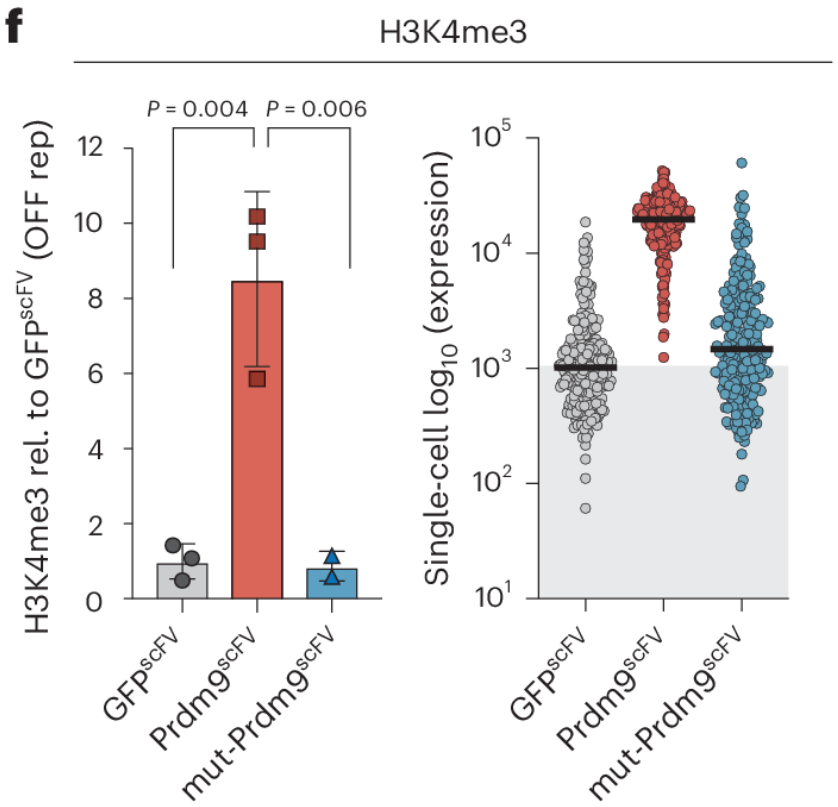
Pengelly…Müller (2013)

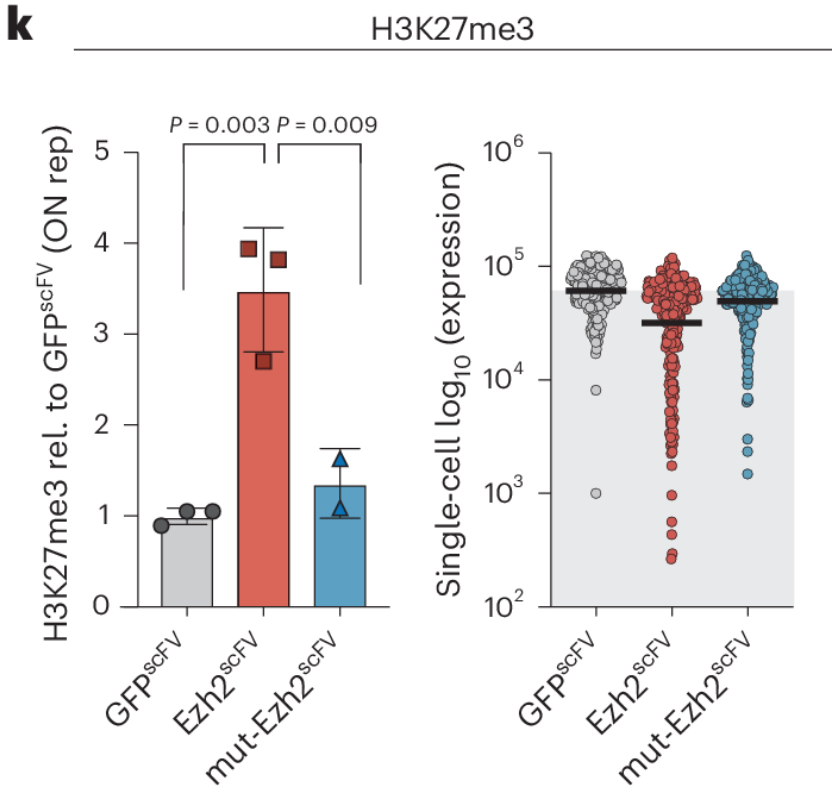


Hödl and Basler (2012)

…But epigenome editing suggests dispensable modifications can drive transcription!



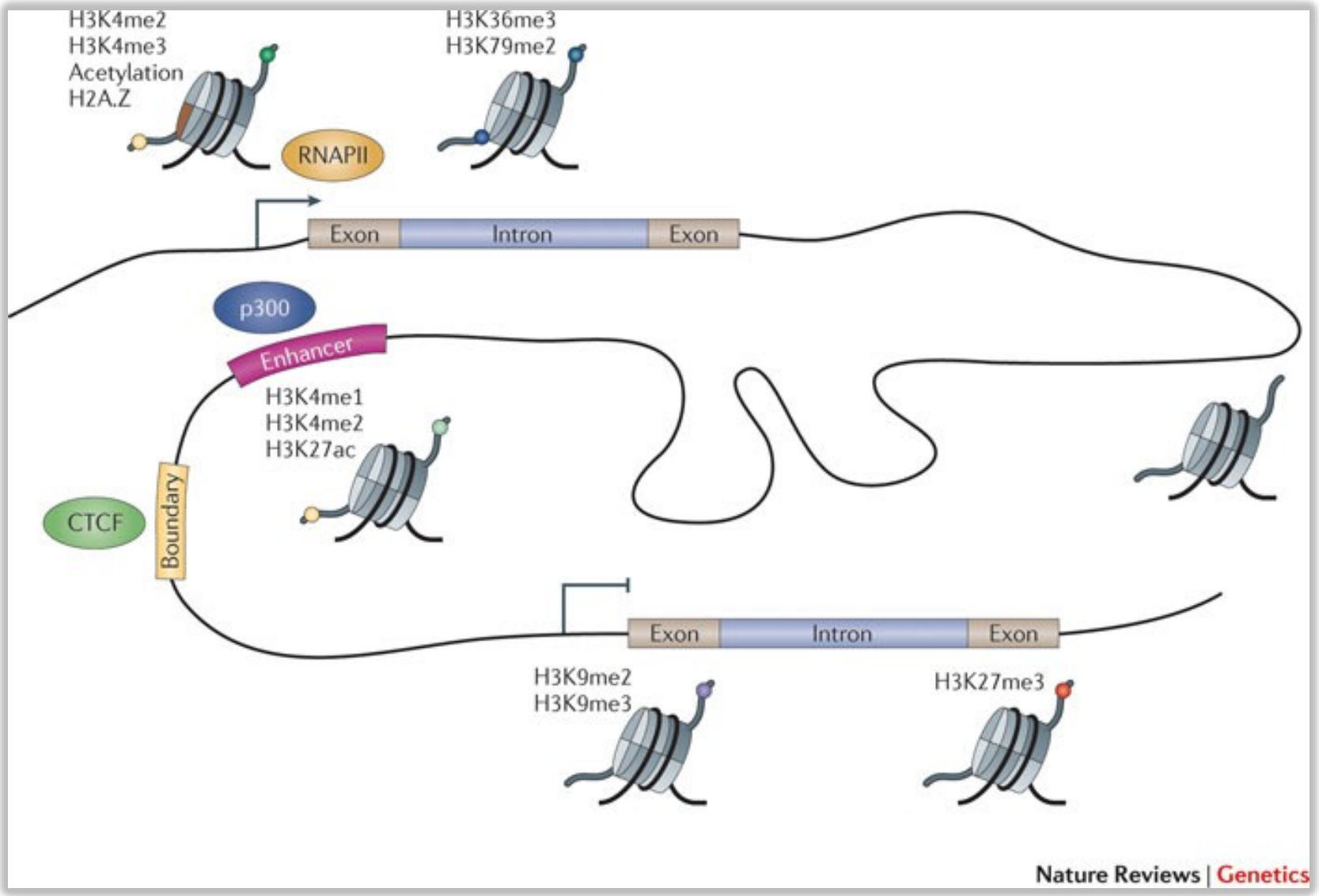




Policarpi…Hackett (2024)

Epigenome provides an annotation of the genome

Chromatin modifications demarcate functional elements in the genome



Zhou, Goren and Bernstein, Nature Rev Genetics, 2011

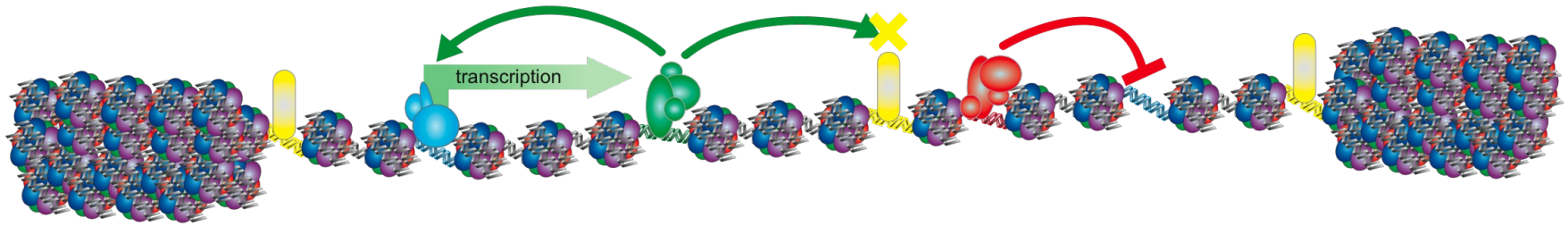
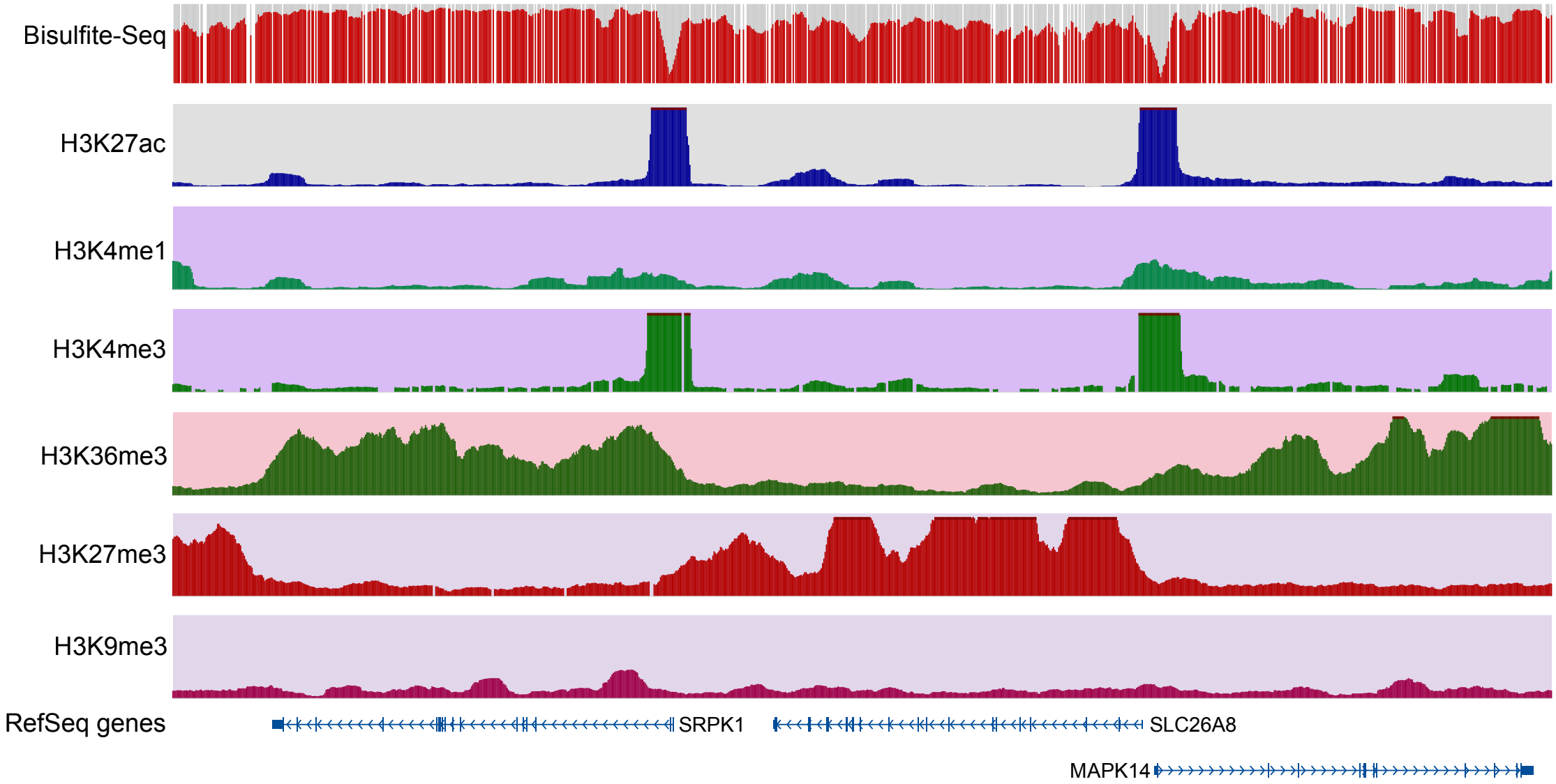


Table 3. Distinctive Chromatin Features of Genomic Elements

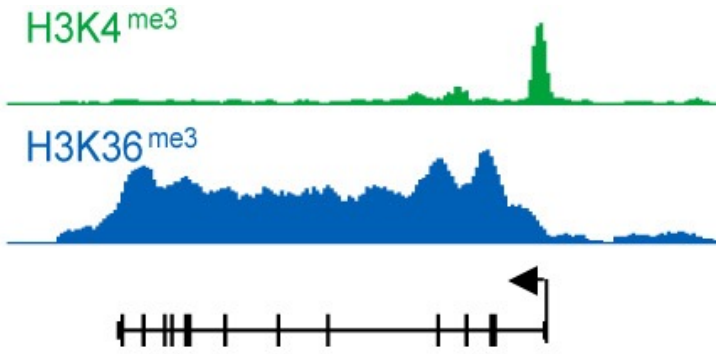
Functional Annotation	Histone Marks
Promoters	H3K4me3
Bivalent/Poised Promoter	H3K4me3/H3K27me3
Transcribed Gene Body	H3K36me3
Enhancer (both active and poised)	H3K4me1
Poised Developmental Enhancer	H3K4me1/H3K27me3
Active Enhancer	H3K4me1/H3K27ac
Polycomb Repressed Regions	H3K27me3
Heterochromatin	H3K9me3

Histone Modifications in Relation to Gene Transcription



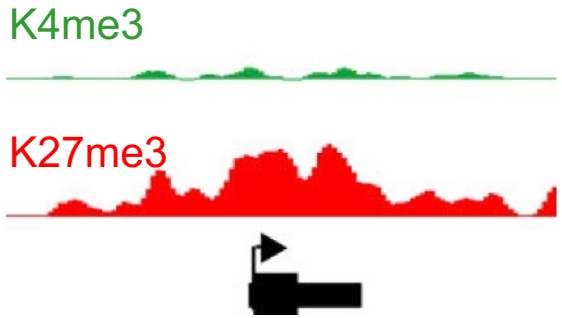
Histone methylation and transcriptional state

Transcribed gene



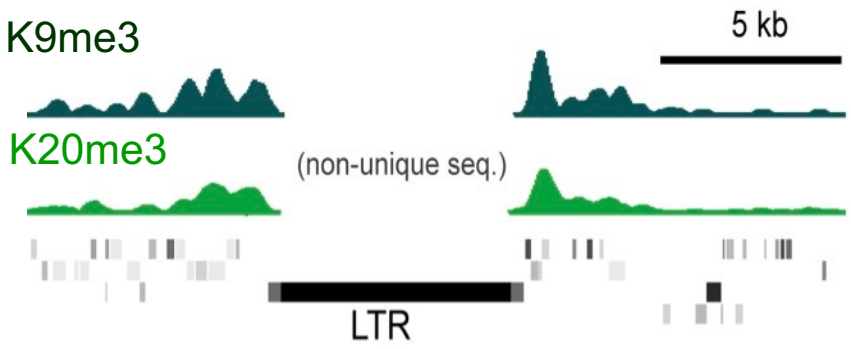
FoxP1

Silent developmental gene

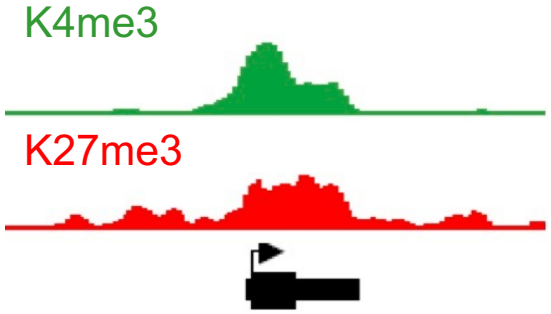


Olig1

Constitutive heterochromatin



'Poised' developmental gene



Olig1

Predicting non-coding RNA?

- From sequence?
  - Not clear which properties can be exploited
  - Sequence features such as promoters are too weak
- Histone modifications + conservation worked

Vol 458|12 March 2009|doi:10.1038/nature07672 nature

LETTERS

Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals

Mitchell Guttman[1,2], Ido Amit[1], Manuel Garber[1], Courtney French[1], Michael F. Lin[1], David Feldser[3], Maite Huarte[1,6], Or Zuk[1], Bryce W. Carey[2,8], John P. Cassady[2,8], Moran N. Cabili[7], Rudolf Jaenisch[2,8], Tarjei S. Mikkelsen[1,4], Tyler Jacks[2,3], Nir Hacohen[1,9], Bradley E. Bernstein[1,10,11], Manolis Kellis[1,5], Aviv Regev[1,2], John L. Rinn[1,6,11]* & Eric S. Lander[1,2,7,8]*

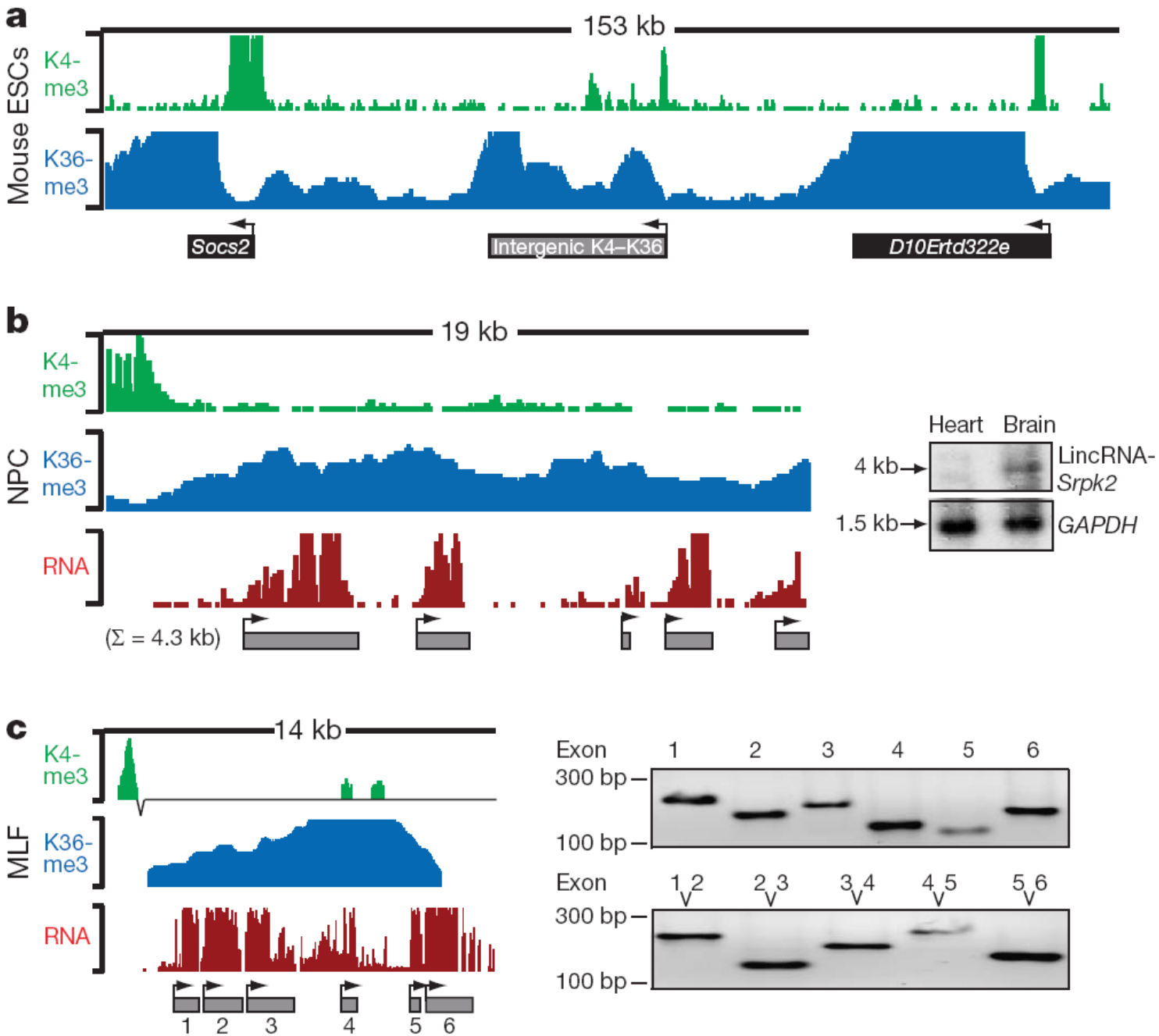
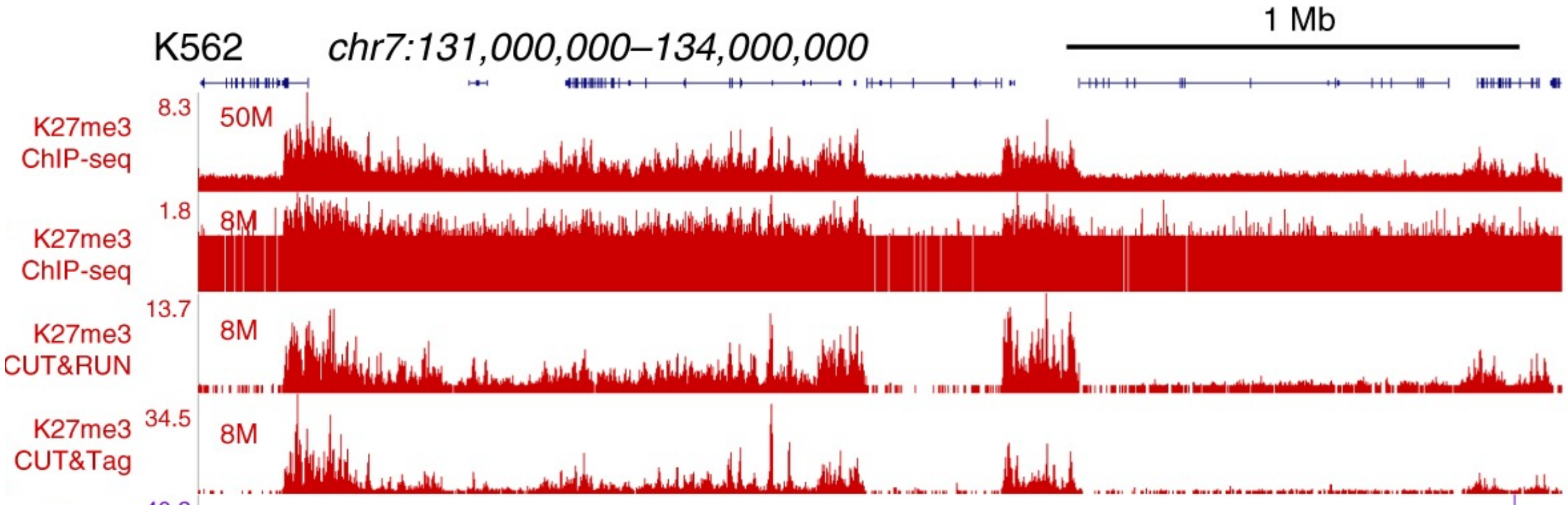


Figure 1 | Intergenic K4–K36 domains produce multi-exonic RNAs.

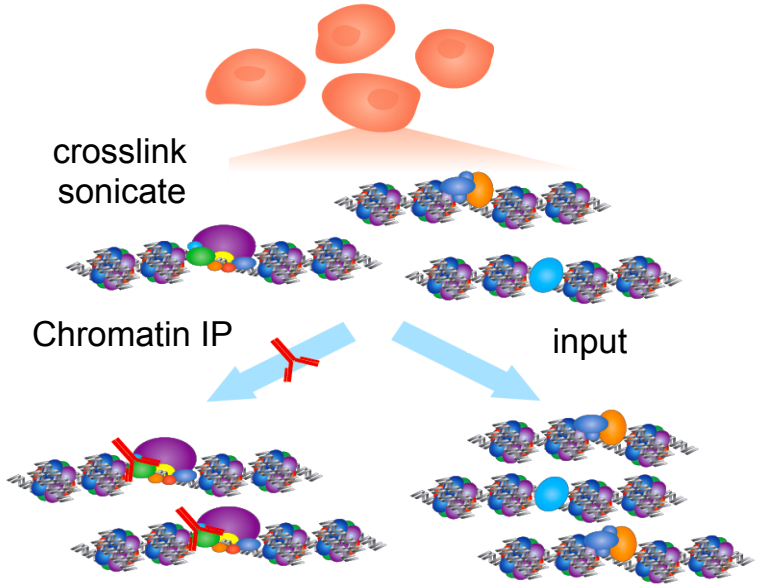
Epigenome profiling: Nucleosome/Histone PTM occupancy



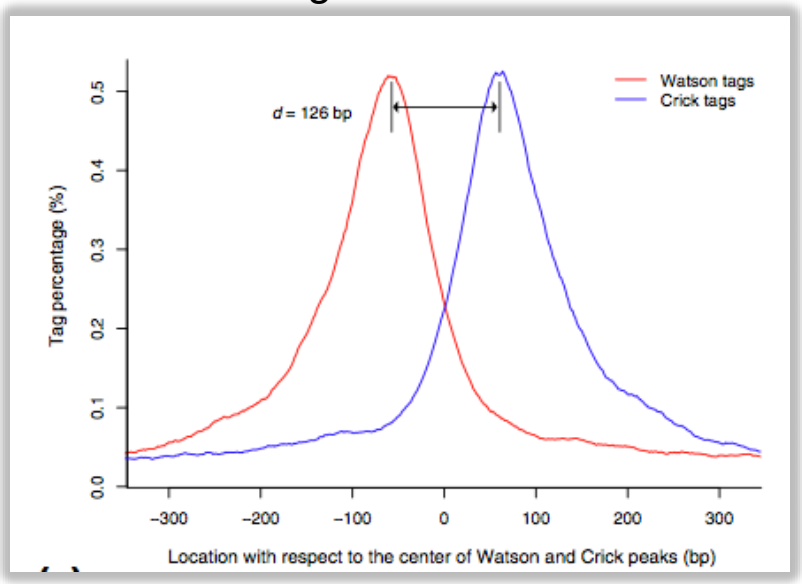
Kaya-Okur et al. (2019)

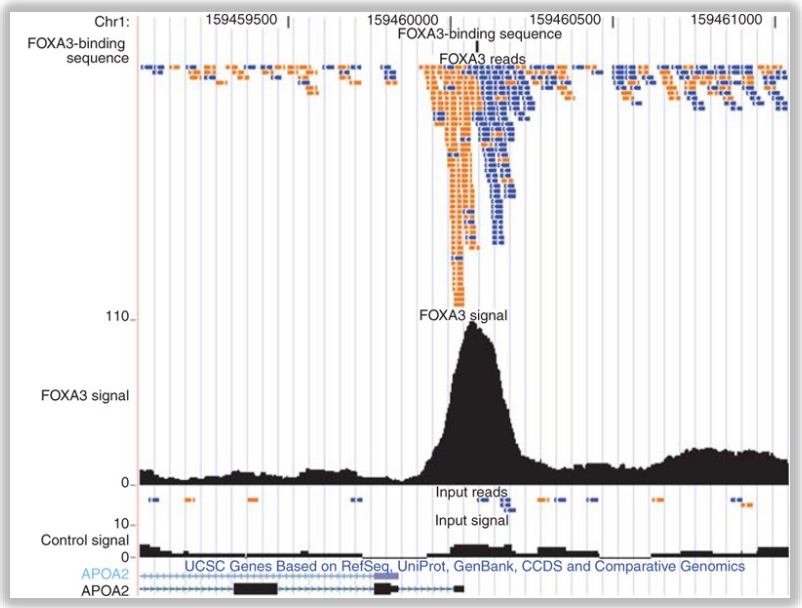
Technologies for Interrogating Chromatin States

ChIP-seq

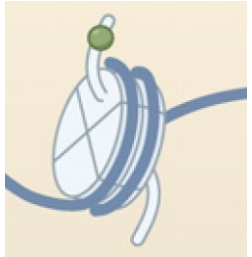


Zhang et al. Genome Biol 2008

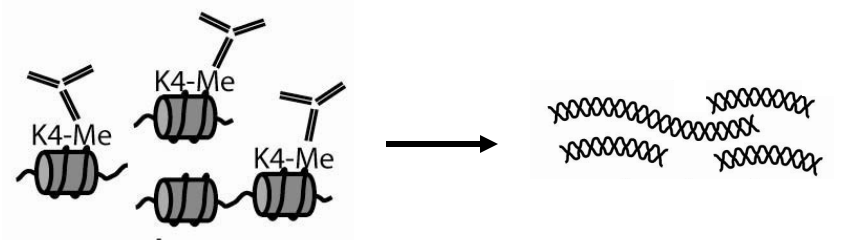




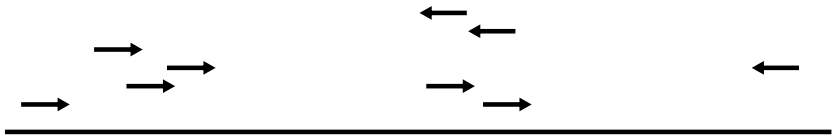
Chromatin-IP Sequencing



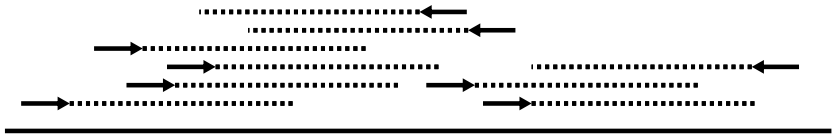
K4me1
K4me2
K4me3 } "active"
K27me3 "repressive"



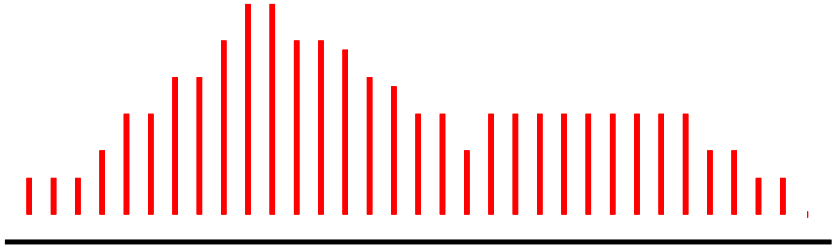
A. Align reads

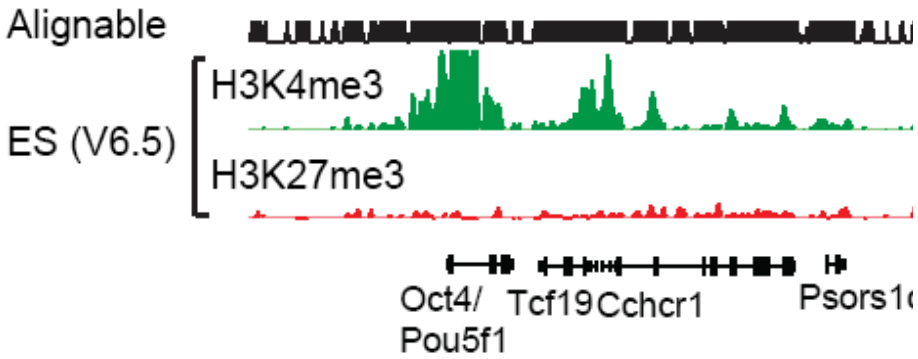


B. Infer positions of ChIP fragments

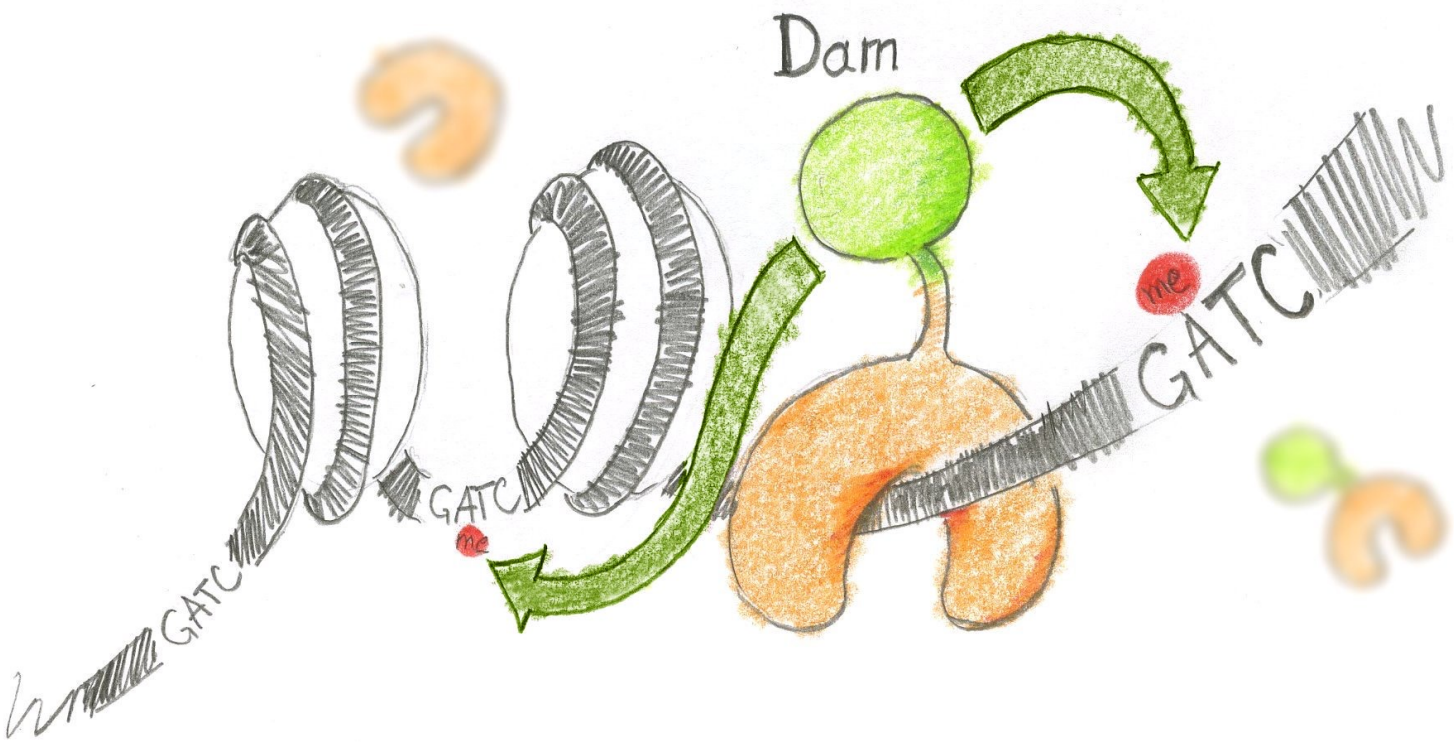


C. Count fragments at each genomic position

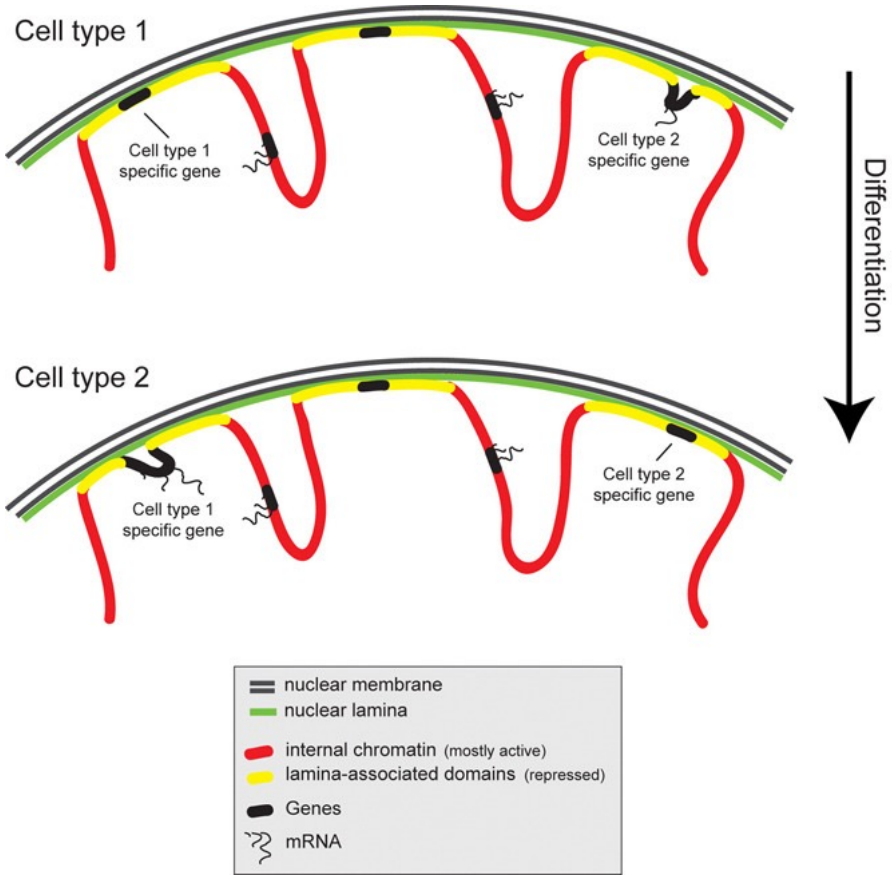




Dam-ID: Enzymatic tethering to detect chromatin spatial interactions

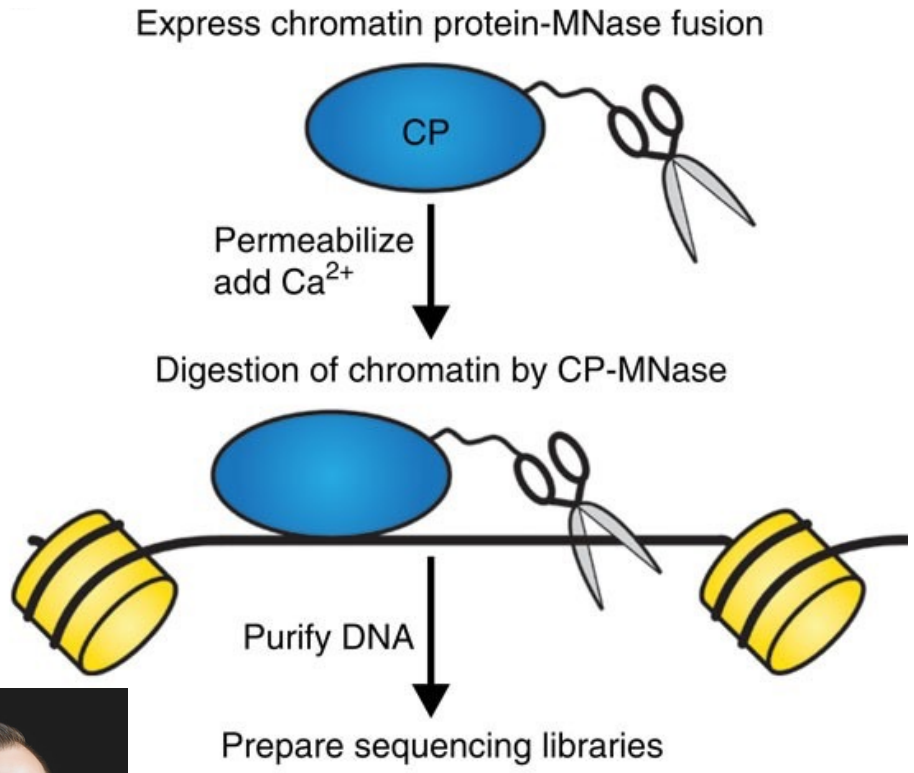


Van Steensel and Henikoff (2000)

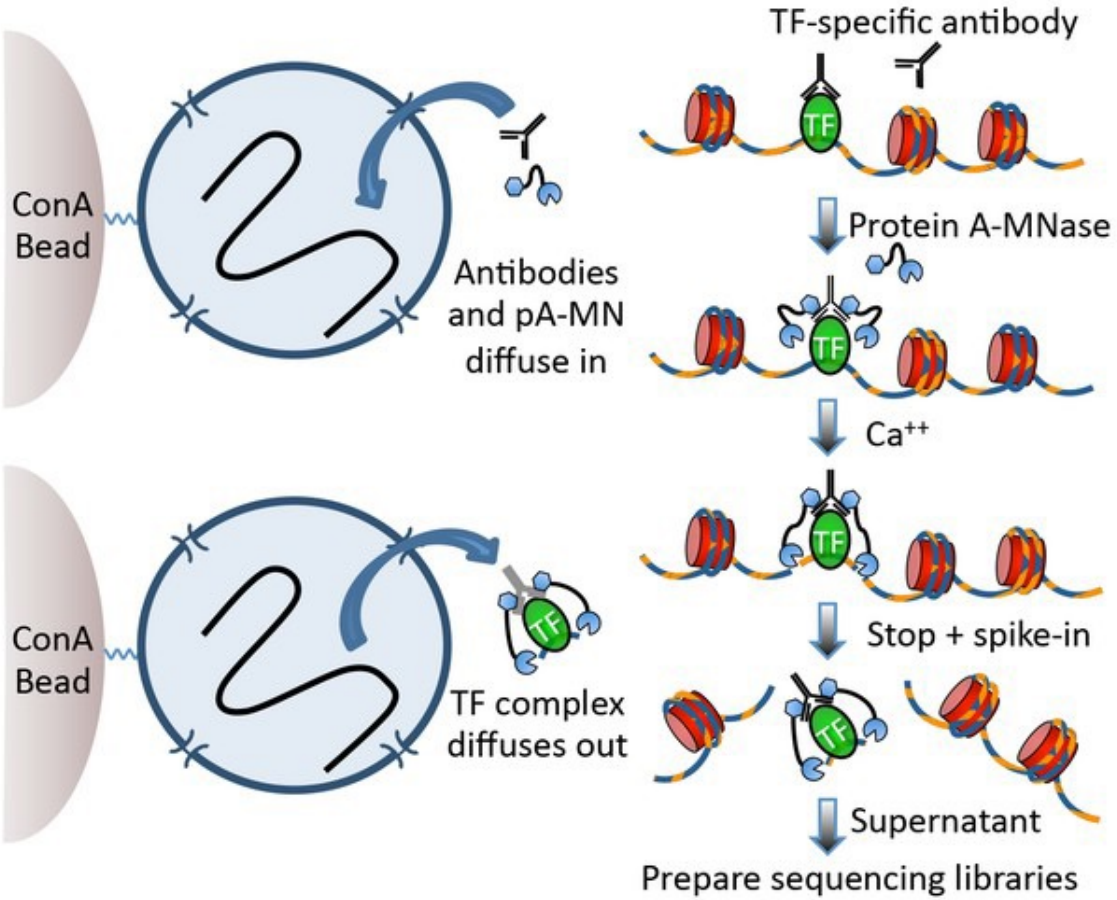


Peric-Hupkes…Van Steensel (2010)

ChEC-seq + CUT&RUN: Tethering nucleases

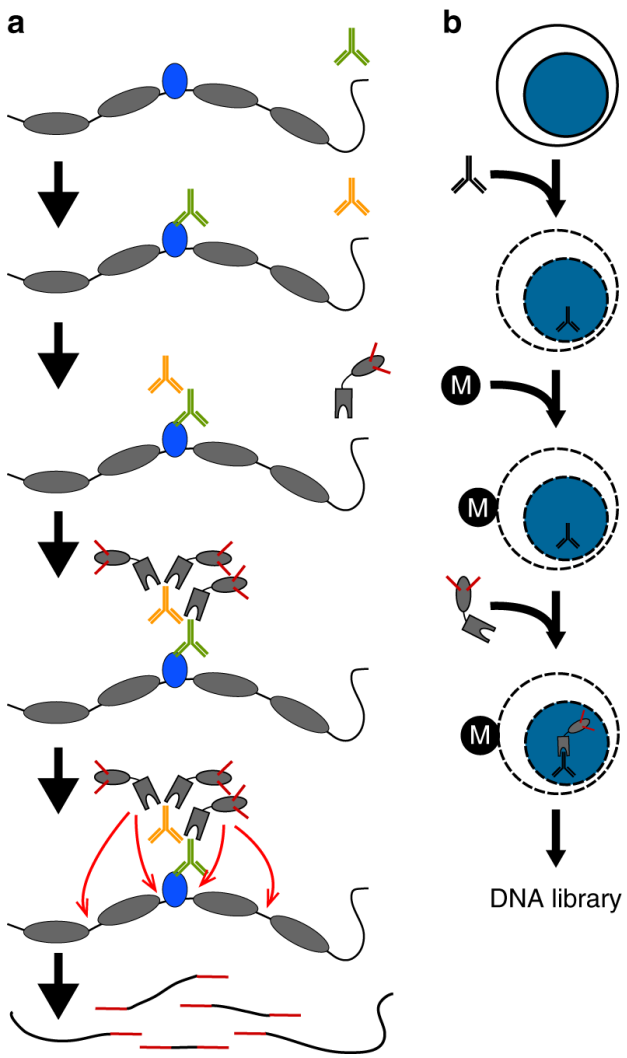


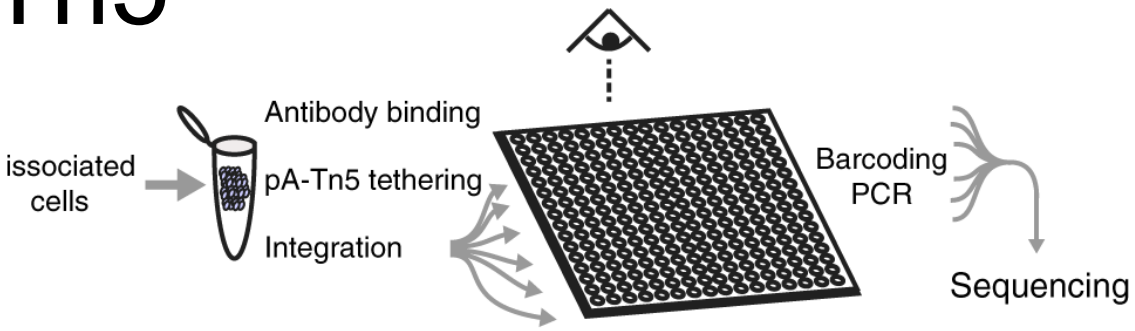
Zentner…Henikoff (2015)

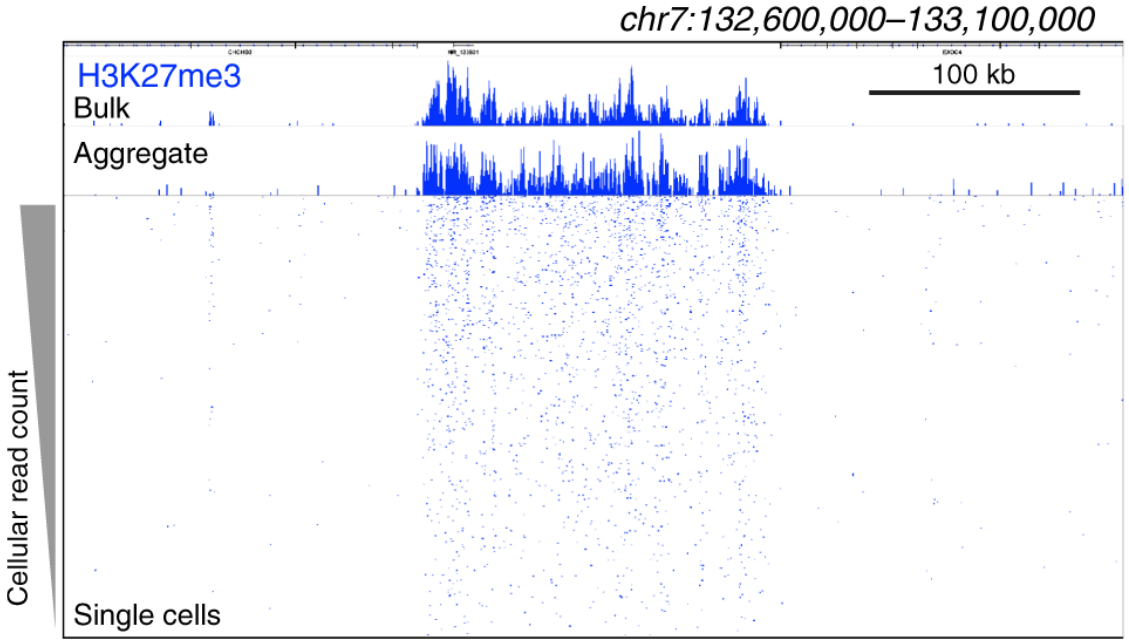


Skene…Henikoff (2017)

CUT&Tag: Tethering Tn5

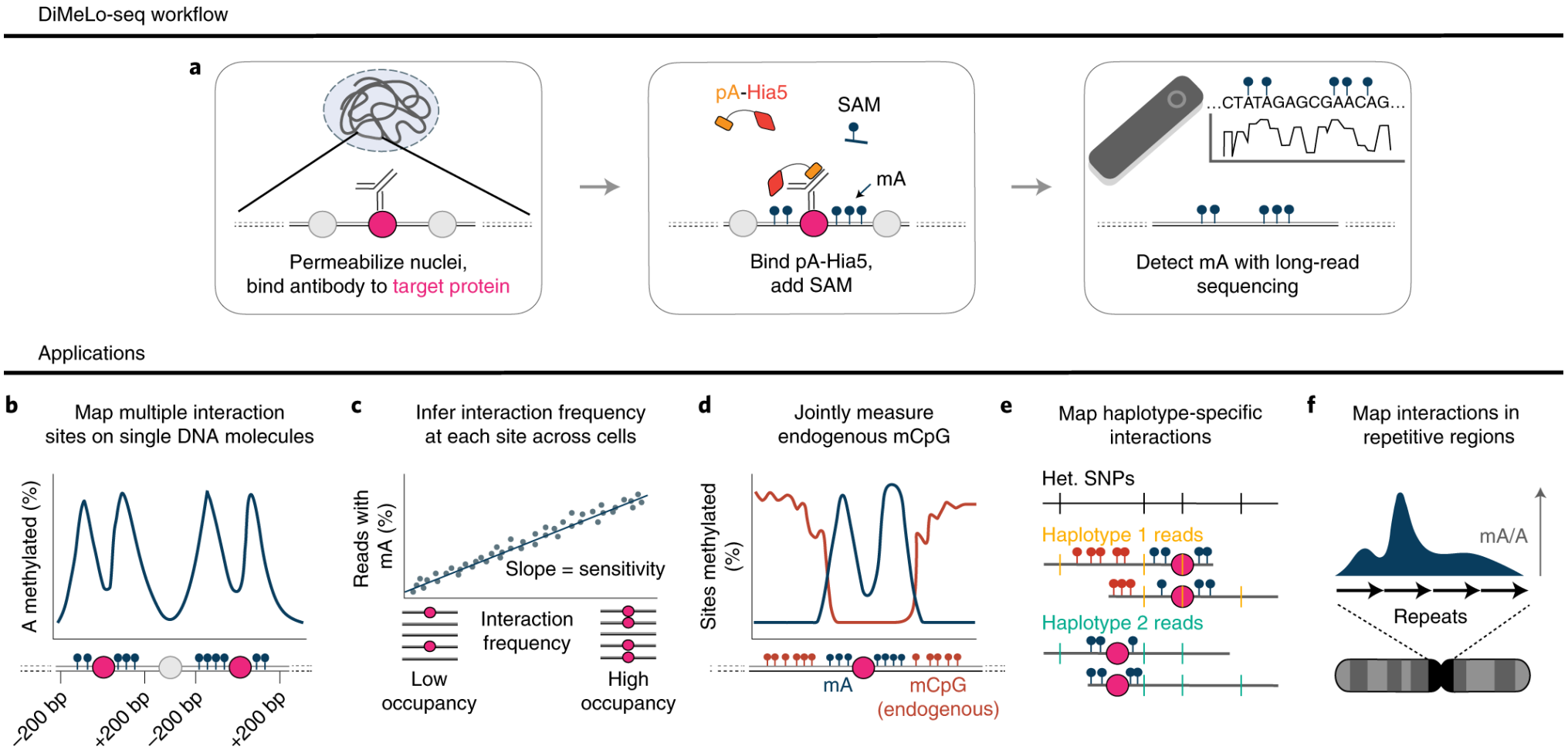






Kaya-Okur…Henikoff (2019)

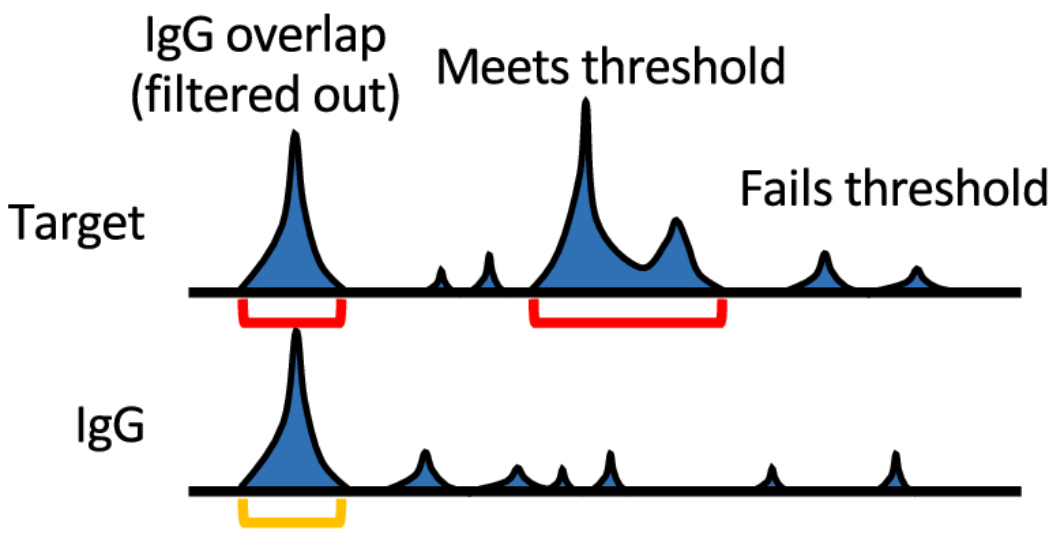
DiMeLo-seq: Tethered methyltransferase + single molecule long-read sequencing



Altemose…Streets (2022)

Analyzing chromatin occupancy data: Peak Calling

- **Feature definition** is a problem for chromatin data: not always defined by gene boundaries (enhancers, silent regions, etc.)
- **Peak calling** identifies regions of the genome that are statistically enriched for sequencing reads to define features of interest
- Major components:
  - A metric for evaluation (e.g. sequencing depth)
  - A background estimator (statistically derived or experimental control)
  - A thresholding approach (e.g. fold change, false discovery rate, etc.)

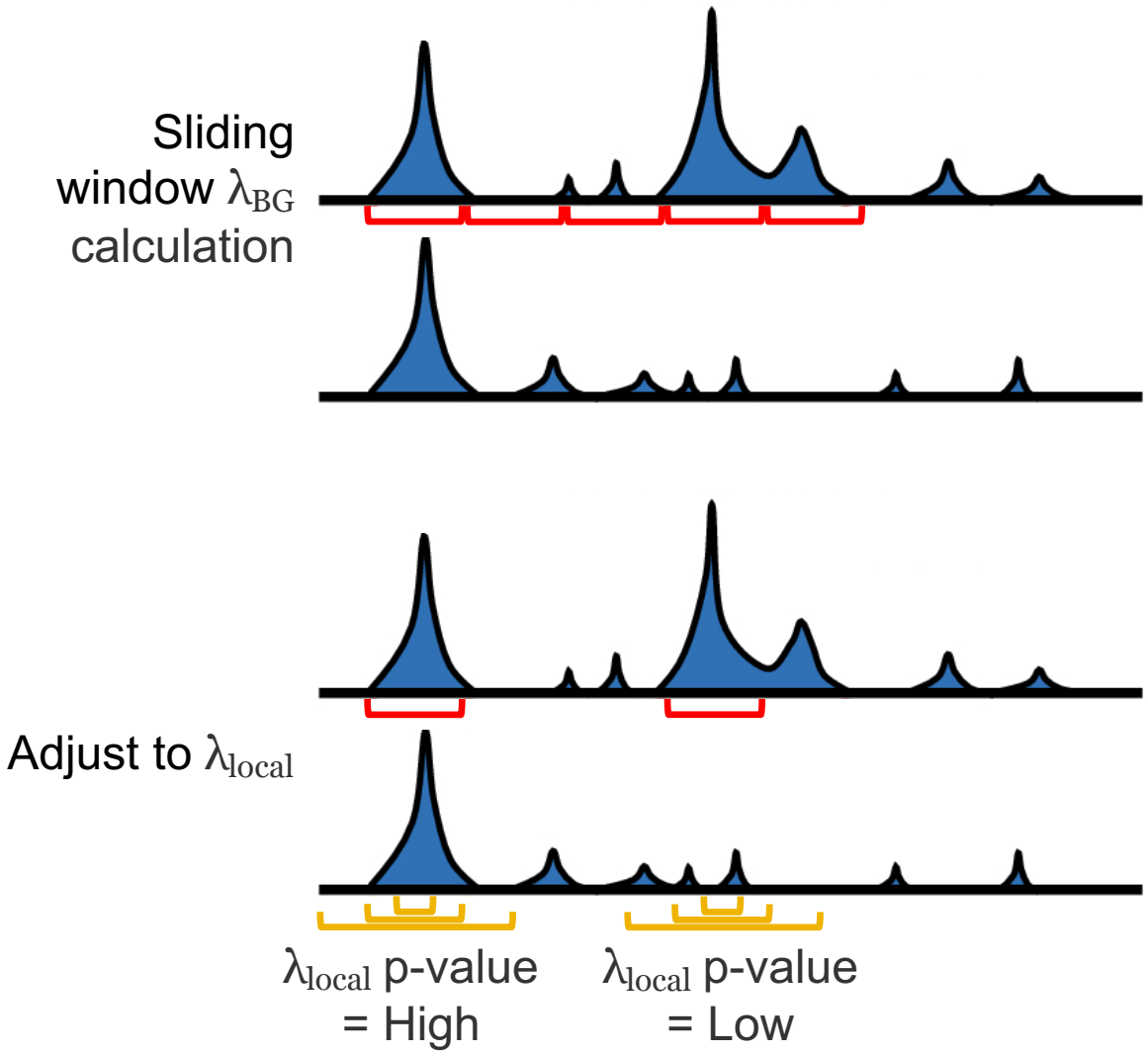


Examples of peak calling approaches

- First ChIP-seq peak caller: Johnson et al. (2007)
  - > 13 sequencing reads within 100 bp
  - < 20% control reads within the same region
- **Arbitrary parameters** can lead to poor performance across different conditions

- Model-based Analysis of ChIP-Seq (MACS) (Zhang 2008)
  - Assumed Poisson-distributed read counts modeled by parameter λ_{BG}
  - Calculate poisson p-value in sliding window of length 2*(median fragment length)

Examples of peak calling approaches



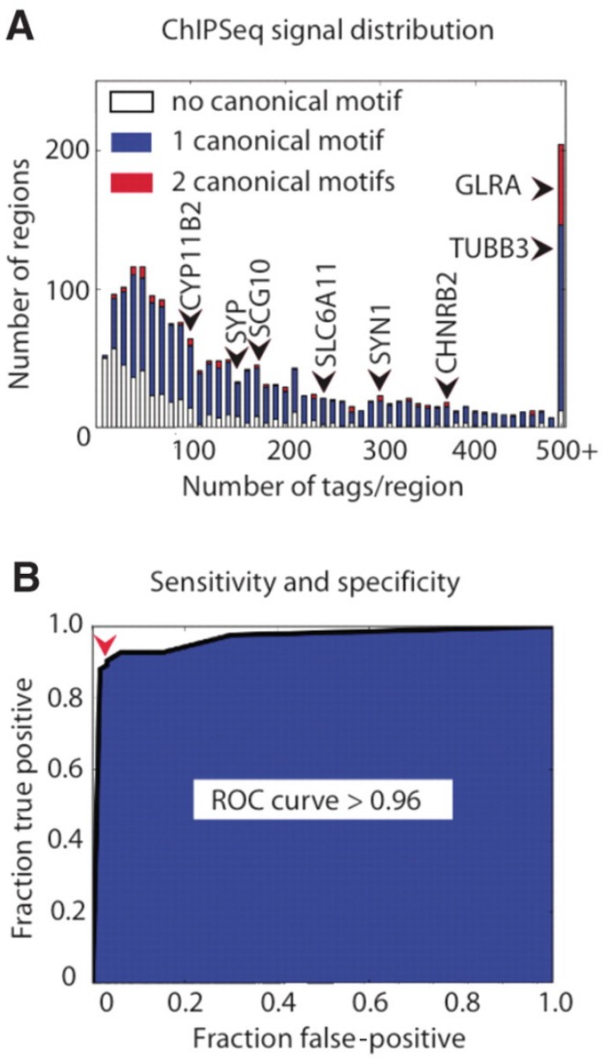
- Model-based Analysis of ChIP-Seq (MACS) (Zhang 2008)
  - Does poisson distribution actually model read distribution properly?
  - Many read depth artifacts (DNA accessibility, library preparation, amplification, etc.)
  - Correct for bias with dynamic $\lambda_{local} = \max(\lambda_{BG}, [\lambda_{1k},] \lambda_{5k}, \lambda_{10k})$

Examples of peak calling approaches

- Other distributions may be more appropriate for non-random read assortment across genome
  - Negative binomial distribution
- How do you set the proper threshold?
  - QC metrics can help to choose the proper FDR cutoff

Quality control for peak calling

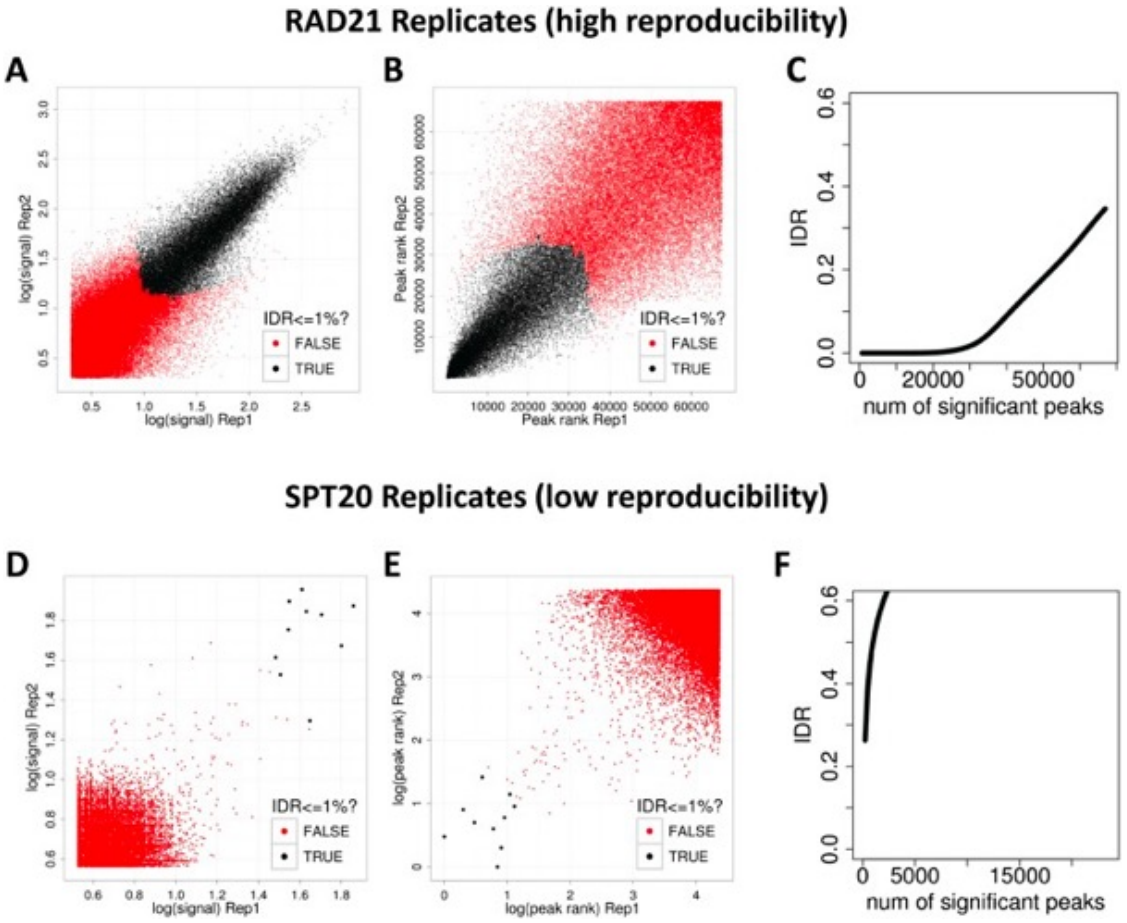
- Biological validation of peak calls
  - AUROC analysis using “gold standard” true positives and negatives
  - Motif enrichment underneath peaks
  - TSS enrichment in peaks (for accessibility, promoter-enriched DNA binding proteins, etc.)



Johnson et al. (2007)

Quality control for peak calling

- Statistical validation of peak calling
  - Peak # vs. read-depth: Call peaks at sub-sampled read depths
  - Fold enrichment vs. read-depth: measure fraction of reads in peaks at different read depths
    - Both of these identify “saturation point” where increased read depth does not increase peak quality
  - Irreproducible Discovery Rate (IDR): Measures variance in ranked lists of peaks called between replicates
    - Ranked by log(signal)
    - Assumes that reproducibility decays as a function of rank



Landt et al. (2012)