

Hi-C and ChIP-seq for Histone modifications, TFBSs, DNAasel Hypersensitivity and ATAC-seq Analysis

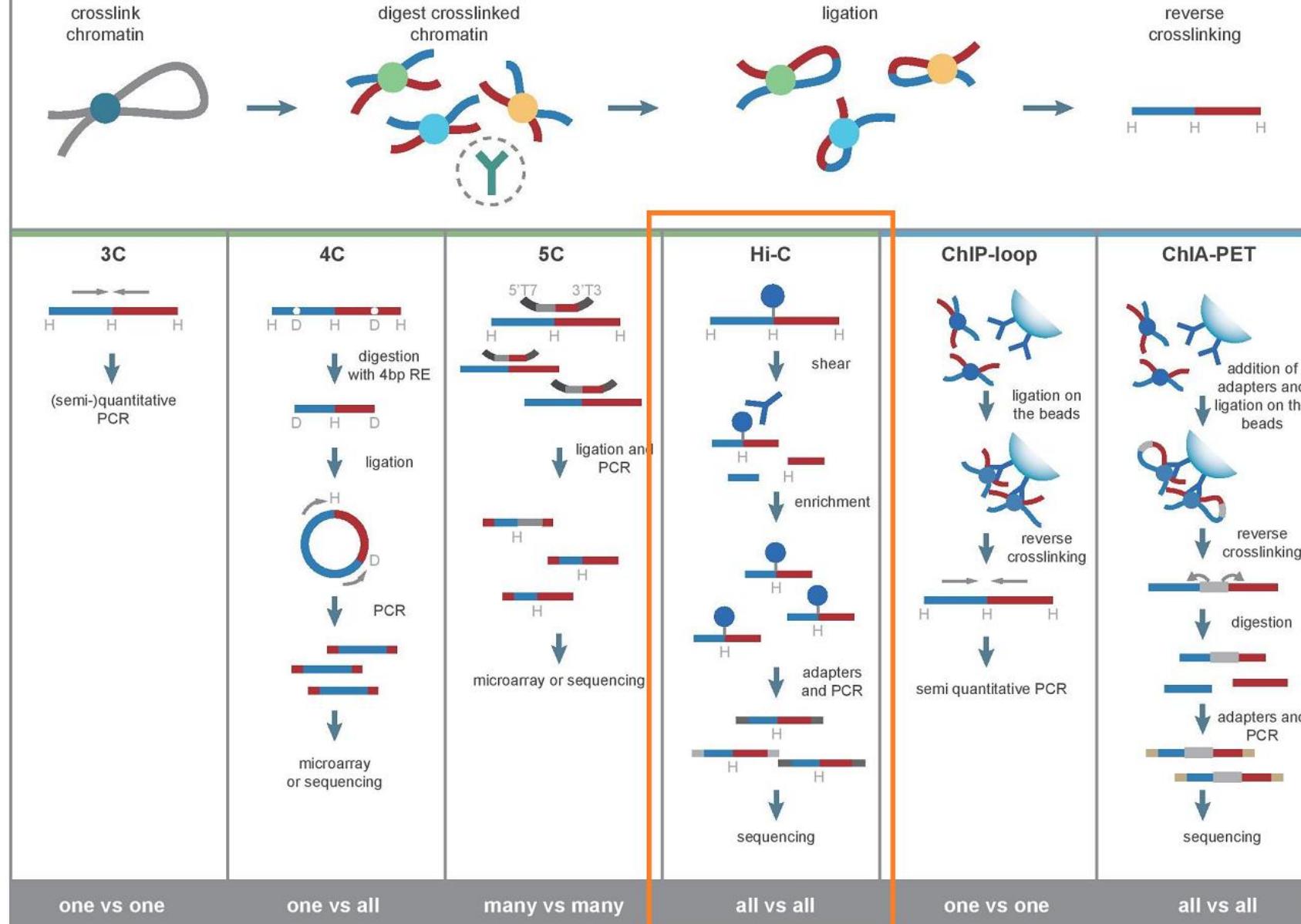
March 29 2025

Phuc-Loi Luu, PhD

Content

- 1. Hi-C**
- 2. ChIP-seq for Histone modifications, TFBSSs, DNAasel Hypersensitivity**
- 3. ATAC-seq**

Chromosome Conformation Technologies





Hi-C

Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome

Erez Lieberman-Aiden *et al.*

Science **326**, 289 (2009);

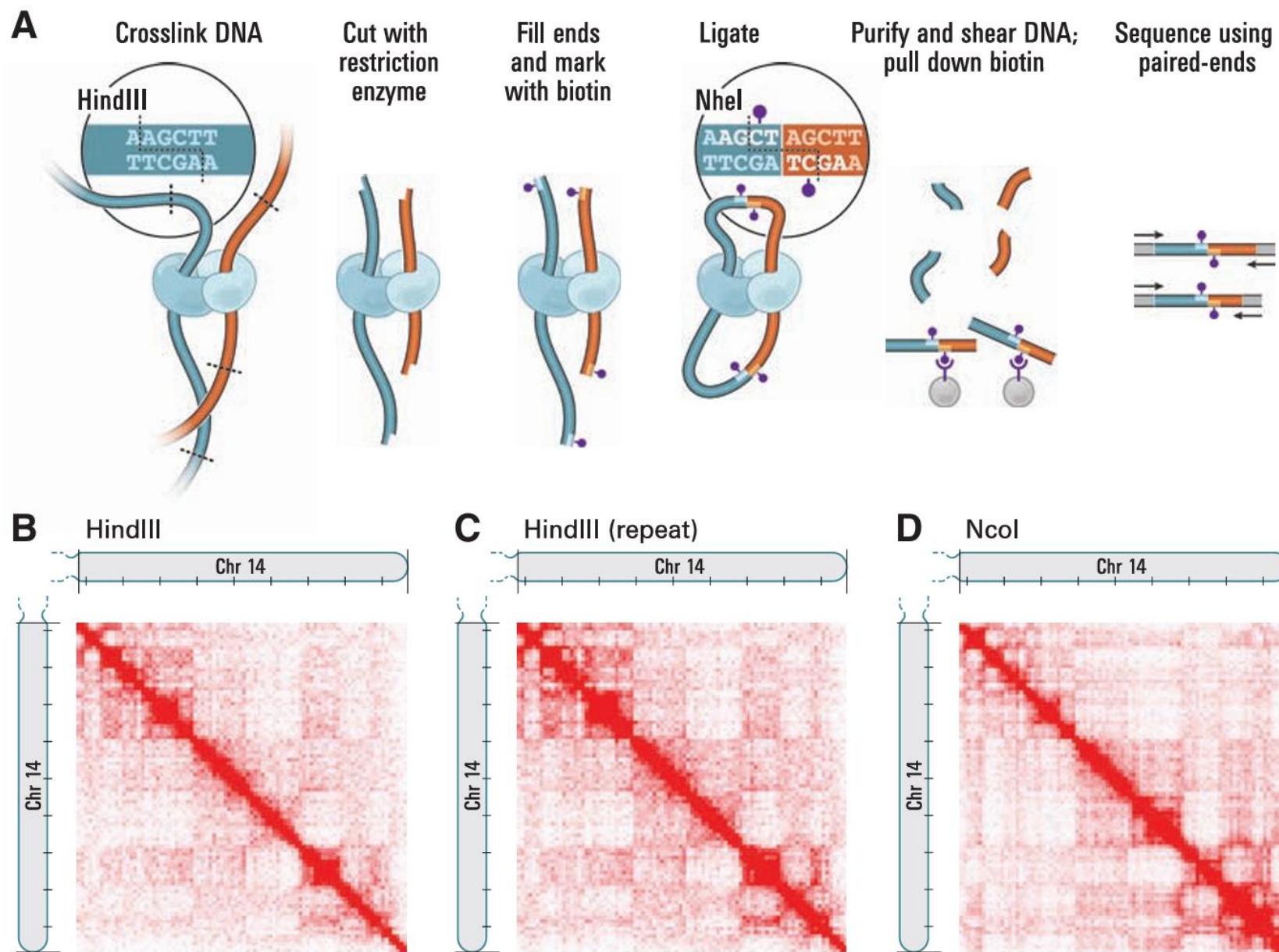
DOI: 10.1126/science.1181369

Erez Lieberman-Aiden,^{1,2,3,4*} Nynke L. van Berkum,^{5*} Louise Williams,¹ Maxim Imakaev,² Tobias Ragoczy,^{6,7} Agnes Telling,^{6,7} Ido Amit,¹ Bryan R. Lajoie,⁵ Peter J. Sabo,⁸ Michael O. Dorschner,⁸ Richard Sandstrom,⁸ Bradley Bernstein,^{1,9} M. A. Bender,¹⁰ Mark Groudine,^{6,7} Andreas Gnirke,¹ John Stamatoyannopoulos,⁸ Leonid A. Mirny,^{2,11} Eric S. Lander,^{1,12,13†} Job Dekker^{5†}

We describe Hi-C, a method that probes the three-dimensional architecture of whole genomes by coupling proximity-based ligation with massively parallel sequencing. We constructed spatial proximity maps of the human genome with Hi-C at a resolution of 1 megabase. These maps confirm the presence of chromosome territories and the spatial proximity of small, gene-rich chromosomes.

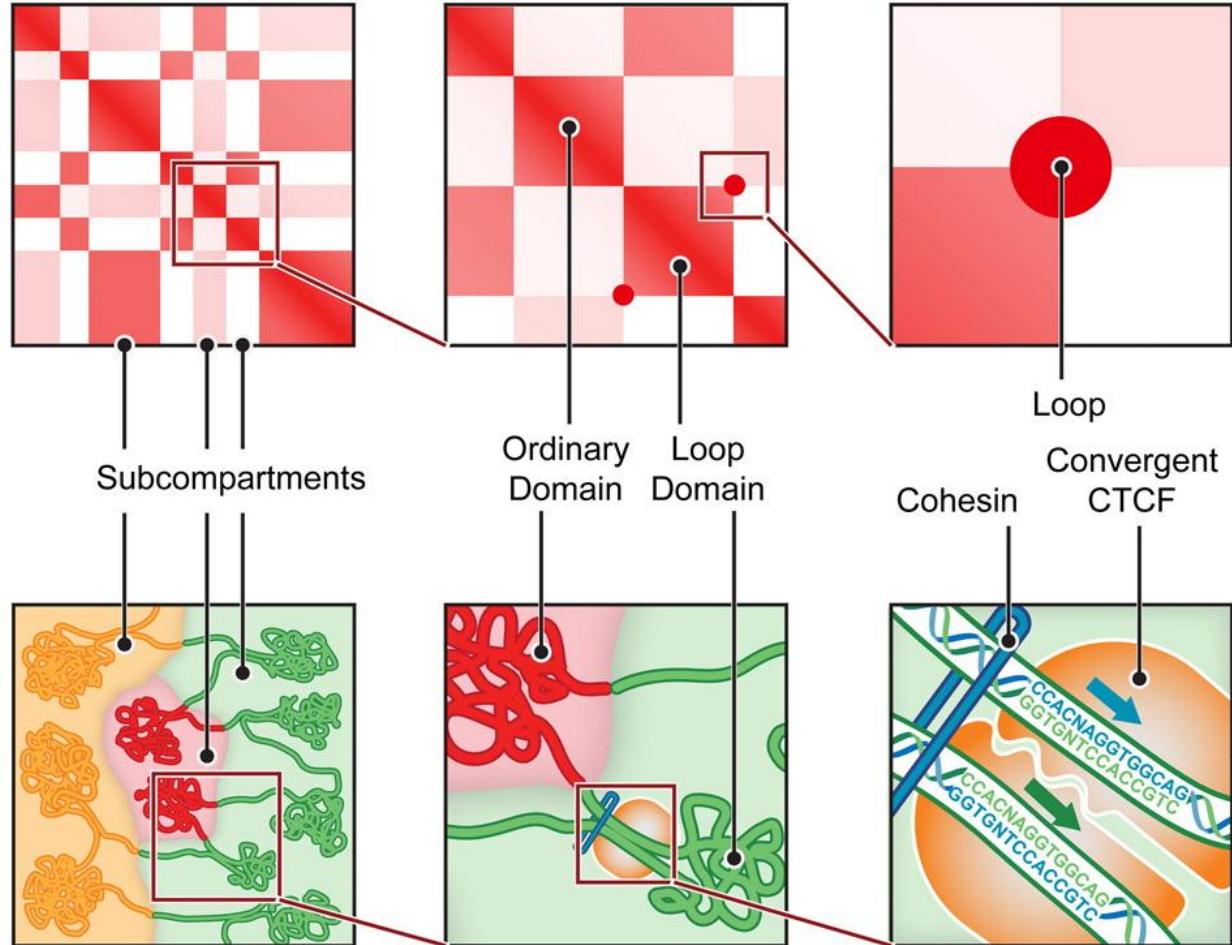
We identified an additional level of genome organization that is characterized by the spatial segregation of open and closed chromatin to form two genome-wide compartments. At the megabase scale, the chromatin conformation is consistent with a fractal globule, a knot-free, polymer conformation that enables maximally dense packing while preserving the ability to easily fold and unfold any genomic locus. The fractal globule is distinct from the more commonly used globular equilibrium model. Our results demonstrate the power of Hi-C to map the dynamic conformations of whole genomes.

Fig. 1. Overview of Hi-C. **(A)** Cells are cross-linked with formaldehyde, resulting in covalent links between spatially adjacent chromatin segments (DNA fragments shown in dark blue, red; proteins, which can mediate such interactions, are shown in light blue and cyan). Chromatin is digested with a restriction enzyme (here, HindIII; restriction site marked by dashed line; see inset), and the resulting sticky ends are filled in with nucleotides, one of which is biotinylated (purple dot). Ligation is performed under extremely dilute conditions to create chimeric molecules; the HindIII site is lost and an NheI site is created (inset). DNA is purified and sheared. Biotinylated junctions are isolated with streptavidin beads and identified by paired-end sequencing. **(B)** Hi-C produces a genome-wide contact matrix. The submatrix shown here corresponds to intrachromosomal interactions on chromosome 14. (Chromosome 14 is acrocentric; the short arm is not shown.) Each pixel represents all interactions between a 1-Mb locus and another 1-Mb locus; intensity corresponds to the total number of reads (0 to 50). Tick marks appear every 10 Mb. **(C and D)** We compared the original experiment with results from a biological repeat using the same restriction enzyme [(C), range from 0 to 50 reads] and with results using a different restriction enzyme [(D), NcoI, range from 0 to 100 reads].



Hi-C Matrices and Models

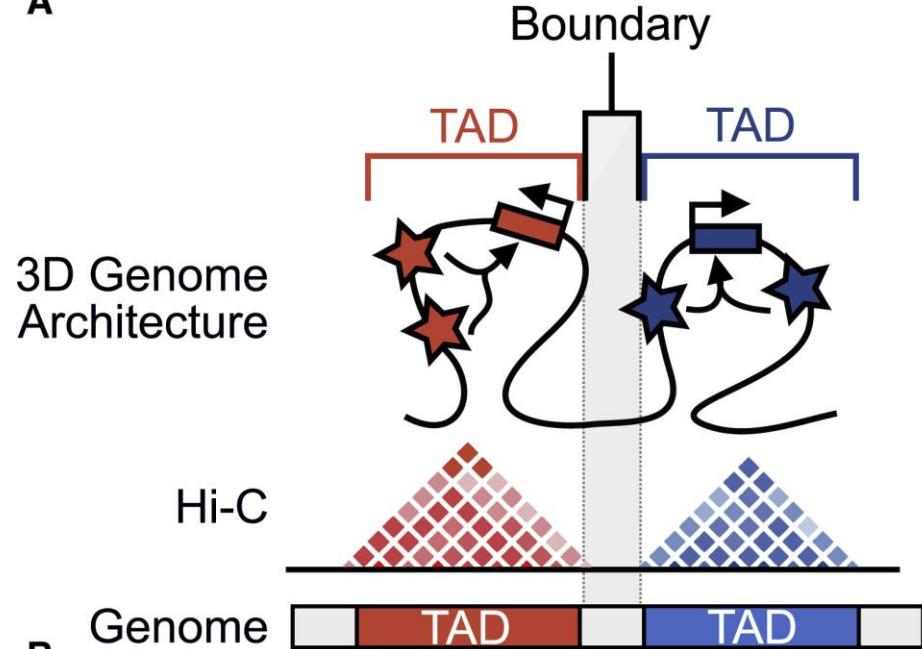
Matrices



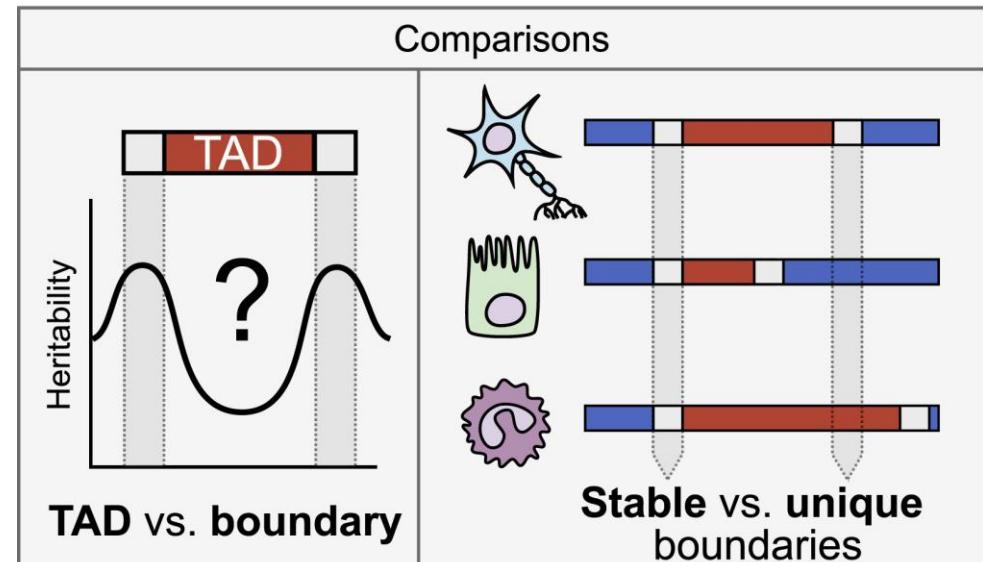
Models

<https://www.cell.com/fulltext/S0092-8674%2814%2901497-4>

A

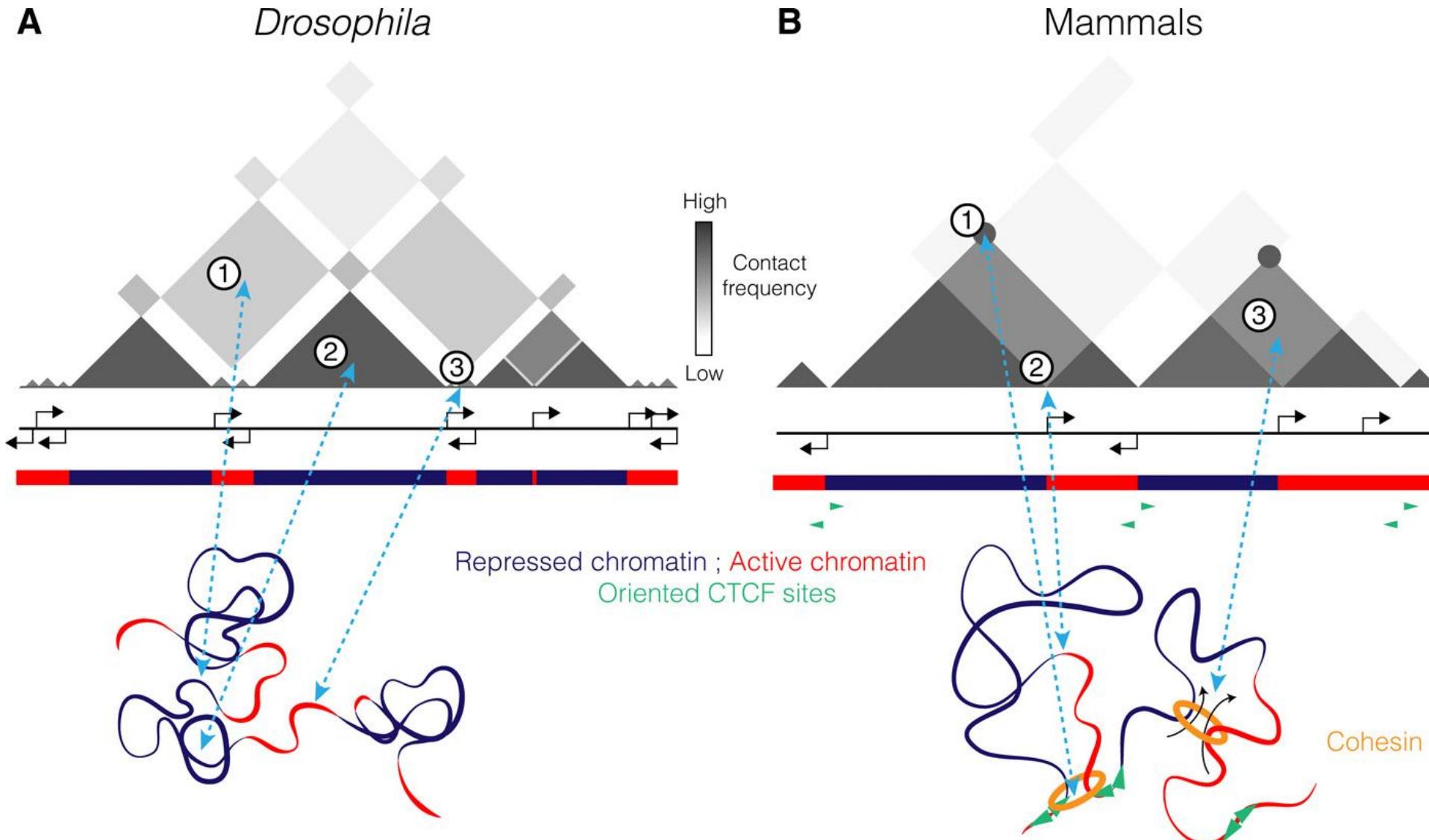


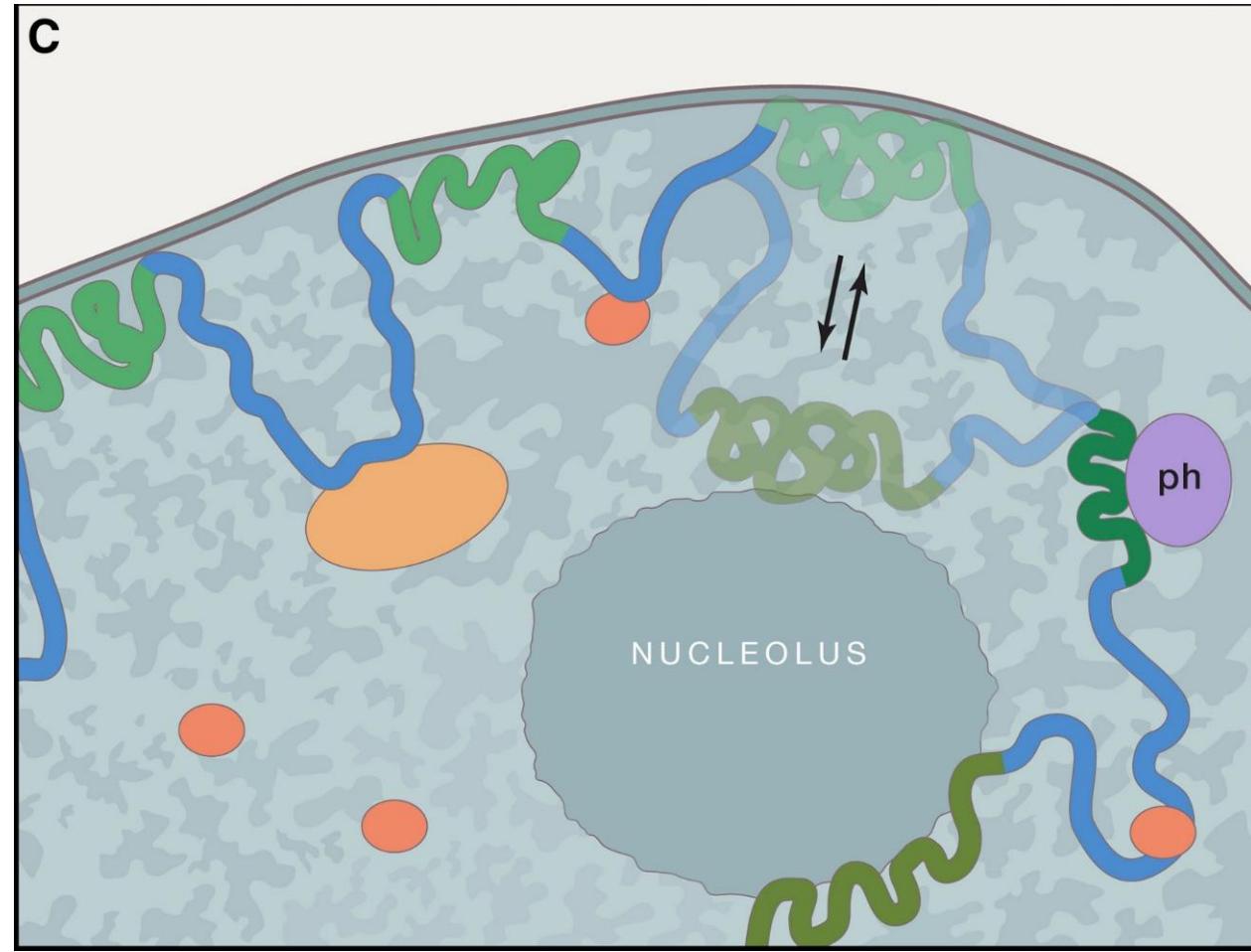
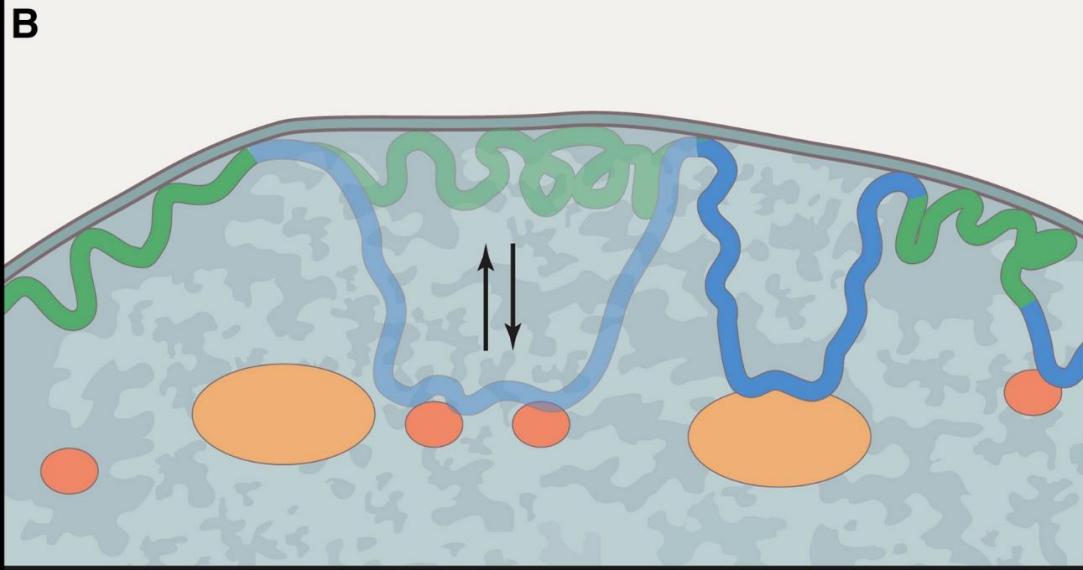
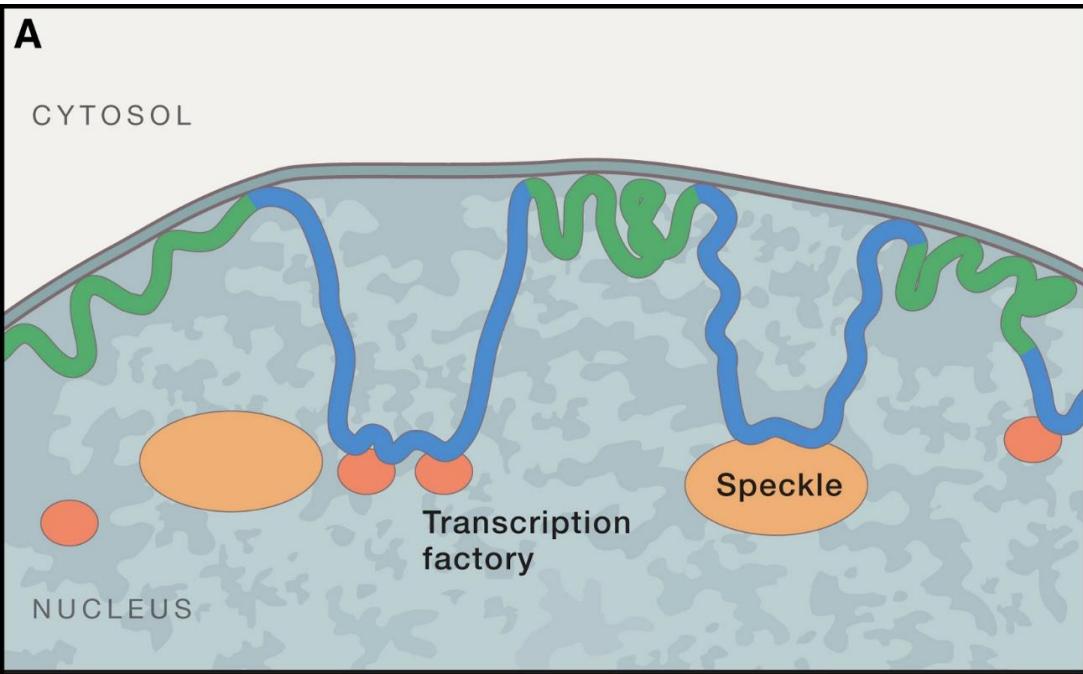
B



<https://www.sciencedirect.com/science/article/pii/S00292972100001X>

Read the Hi-C data (0)



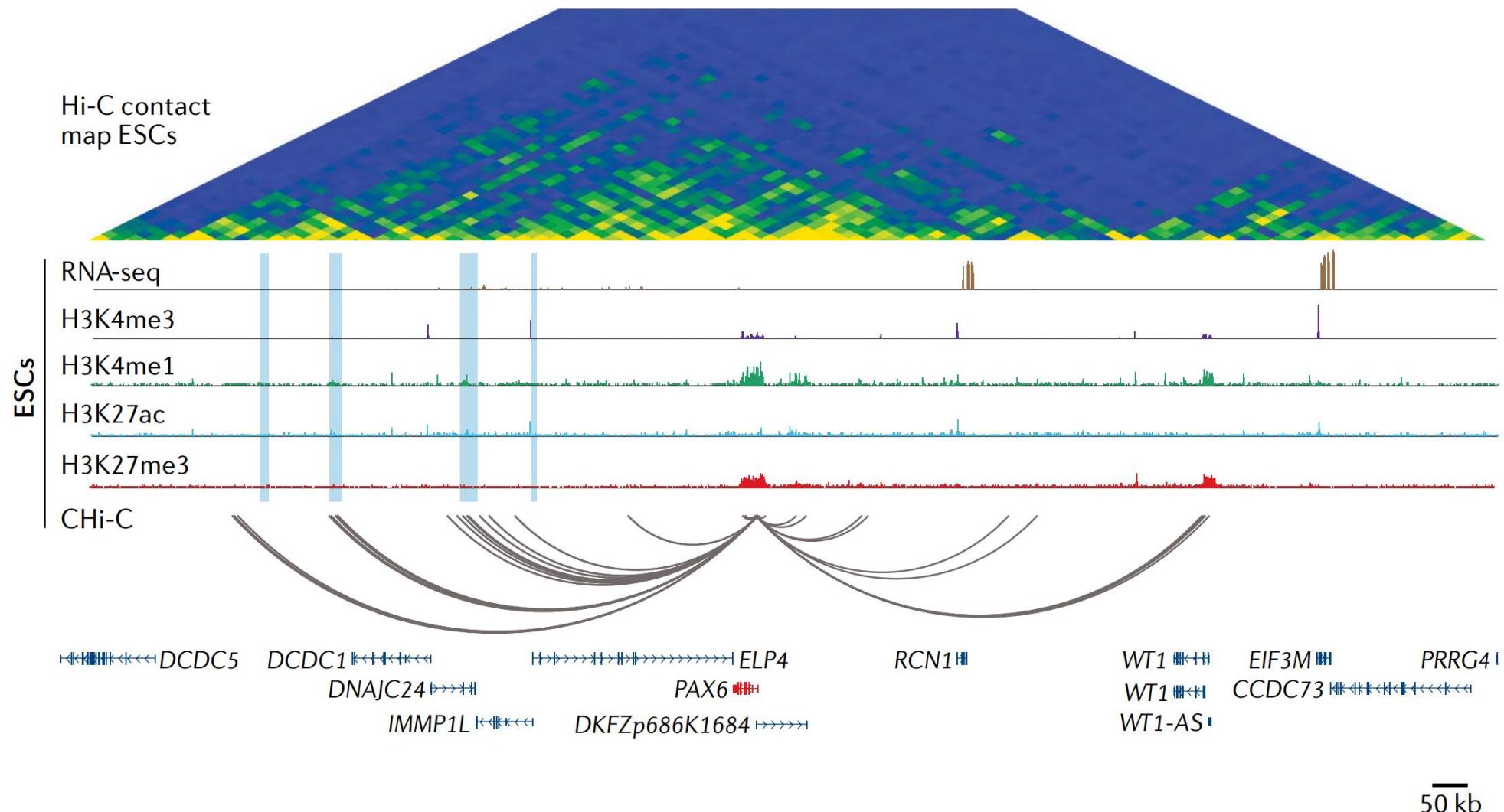


(A) Besides anchoring of LADs (green) to the NL, other regions (blue) may be tethered to nuclear structures that are permissive for transcription (orange), such as transcription factories (tf) or splicing factor speckles (speckles).

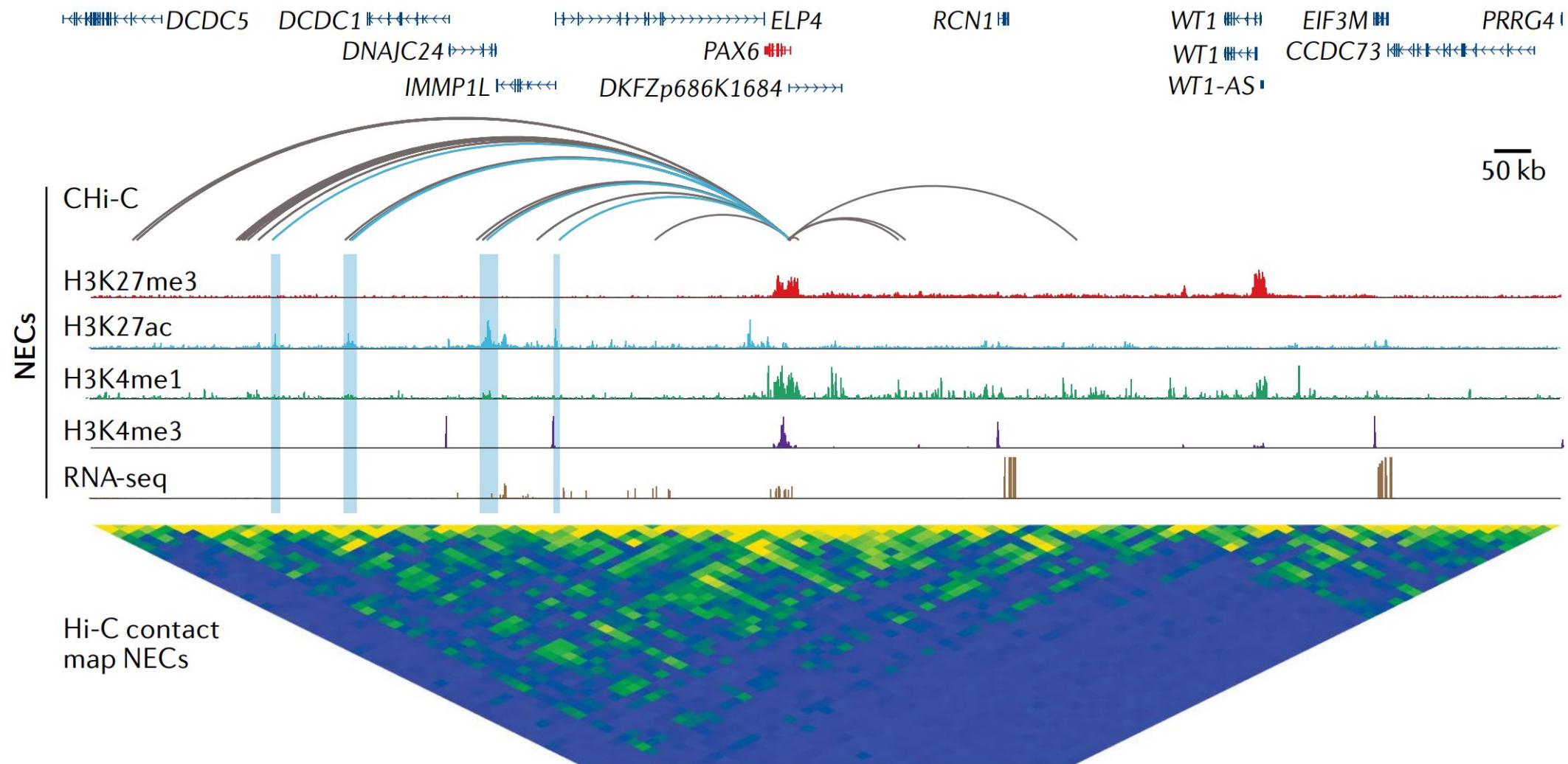
(B) Some LADs (semi-transparent green) contact the NL erratically (i.e., in a subset of cells) and may become transcriptionally active when associated with a permissive compartment (semi-transparent blue).

(C) Some LADs are apparently stochastically distributed between the NL, nucleoli, and pericentromeric heterochromatin (ph), which are all repressive environments.

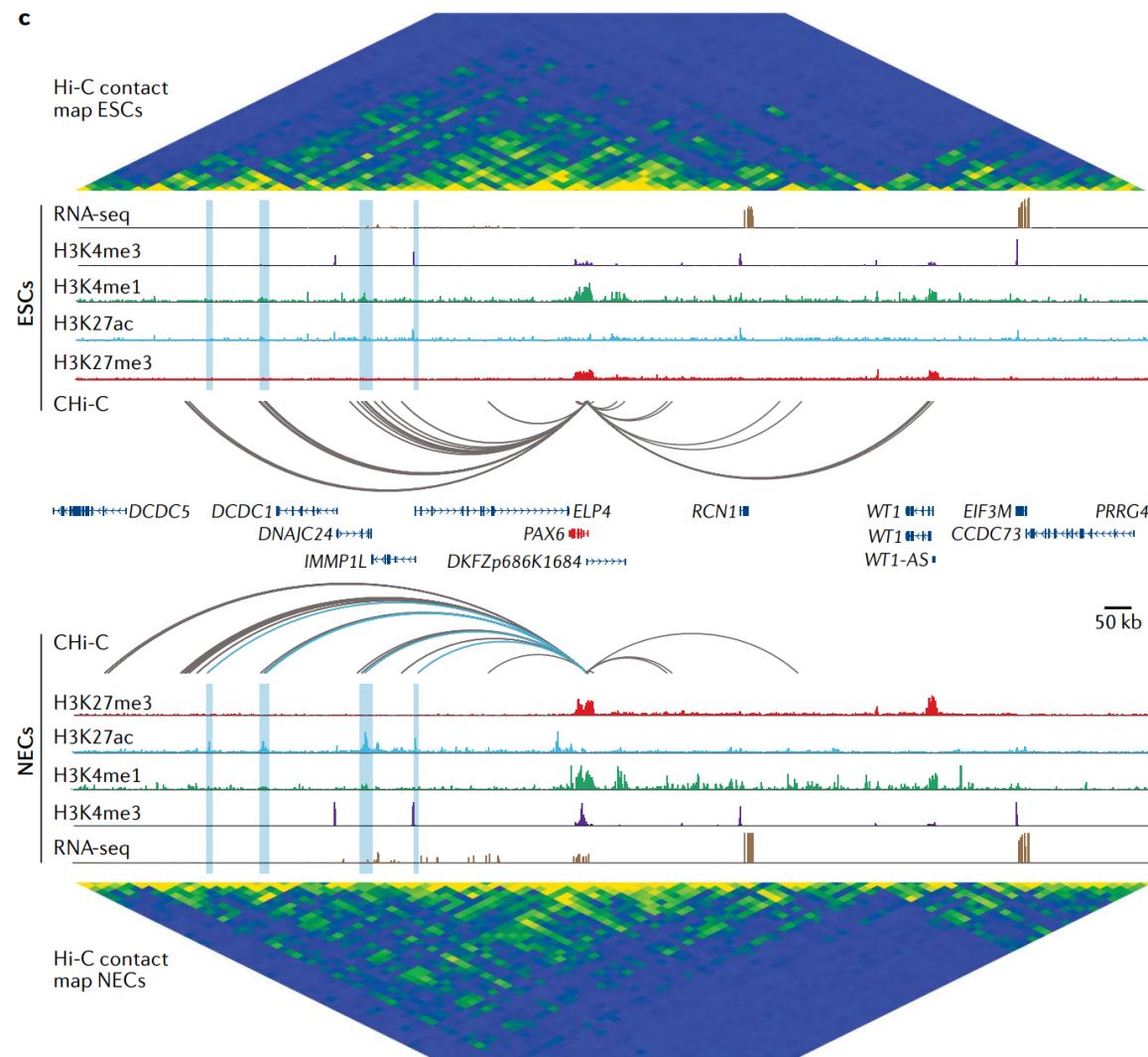
Read the Hi-C data (1)



Read the Hi-C data (2)



Read the Hi-C data (3)



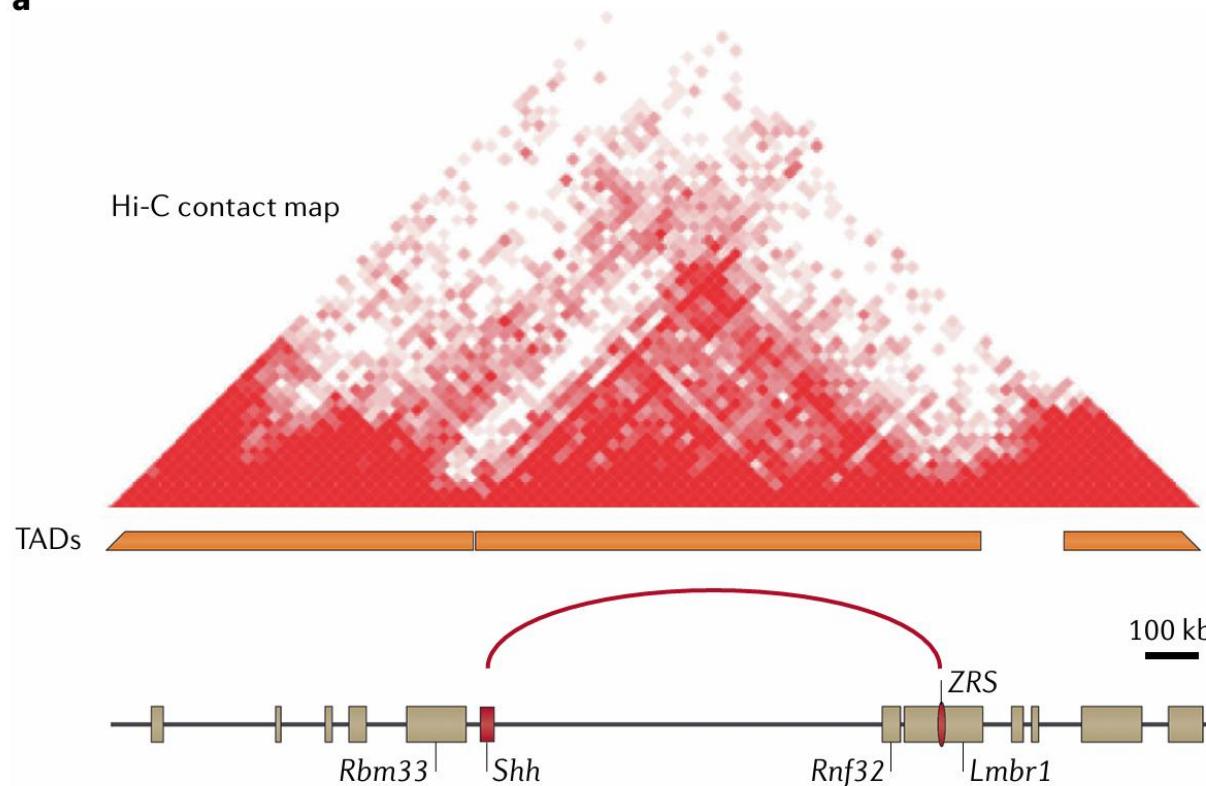
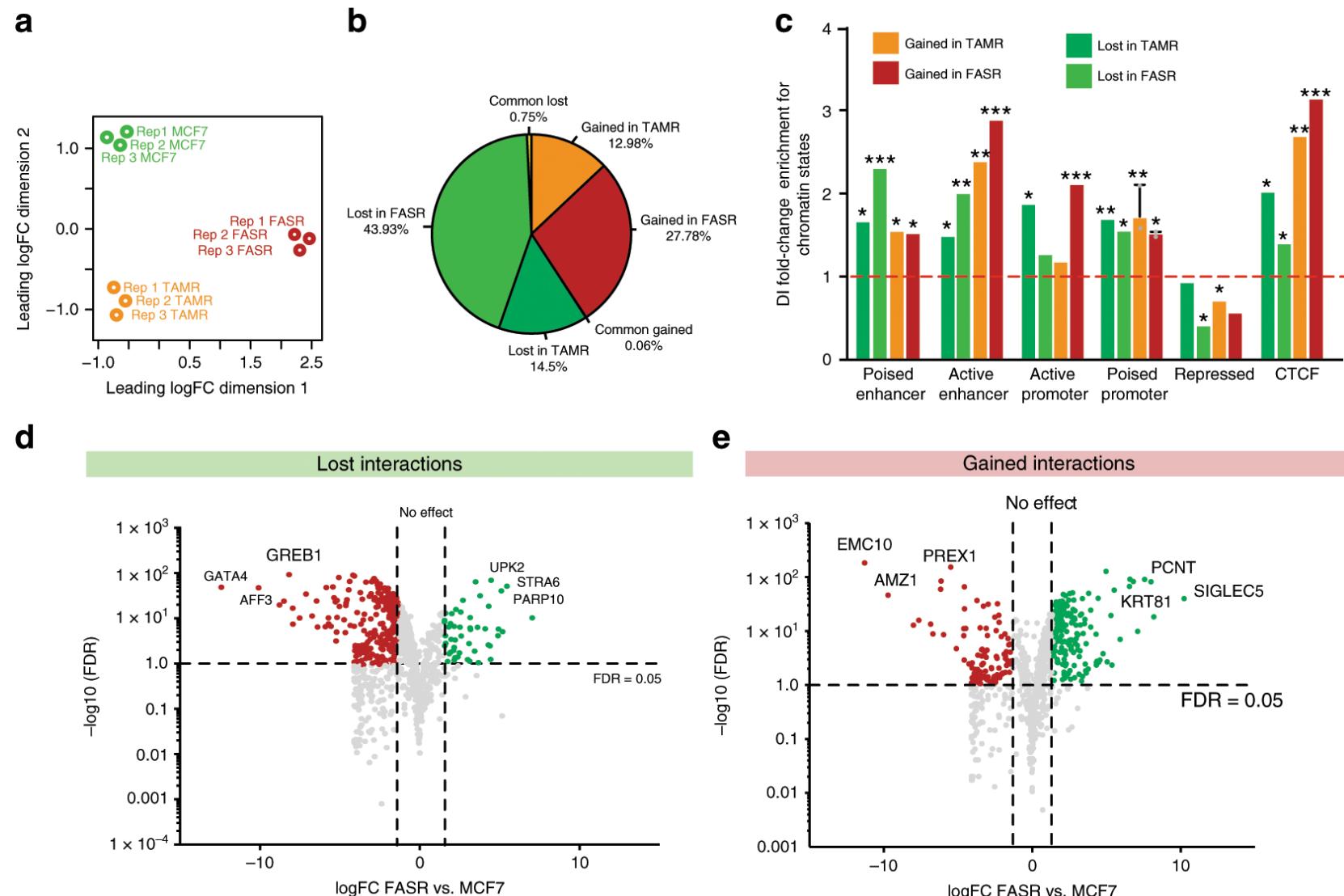
a**b**

Fig. 2 | The Shh limb bud enhancer: a paradigm for long-range enhancer control of gene expression.

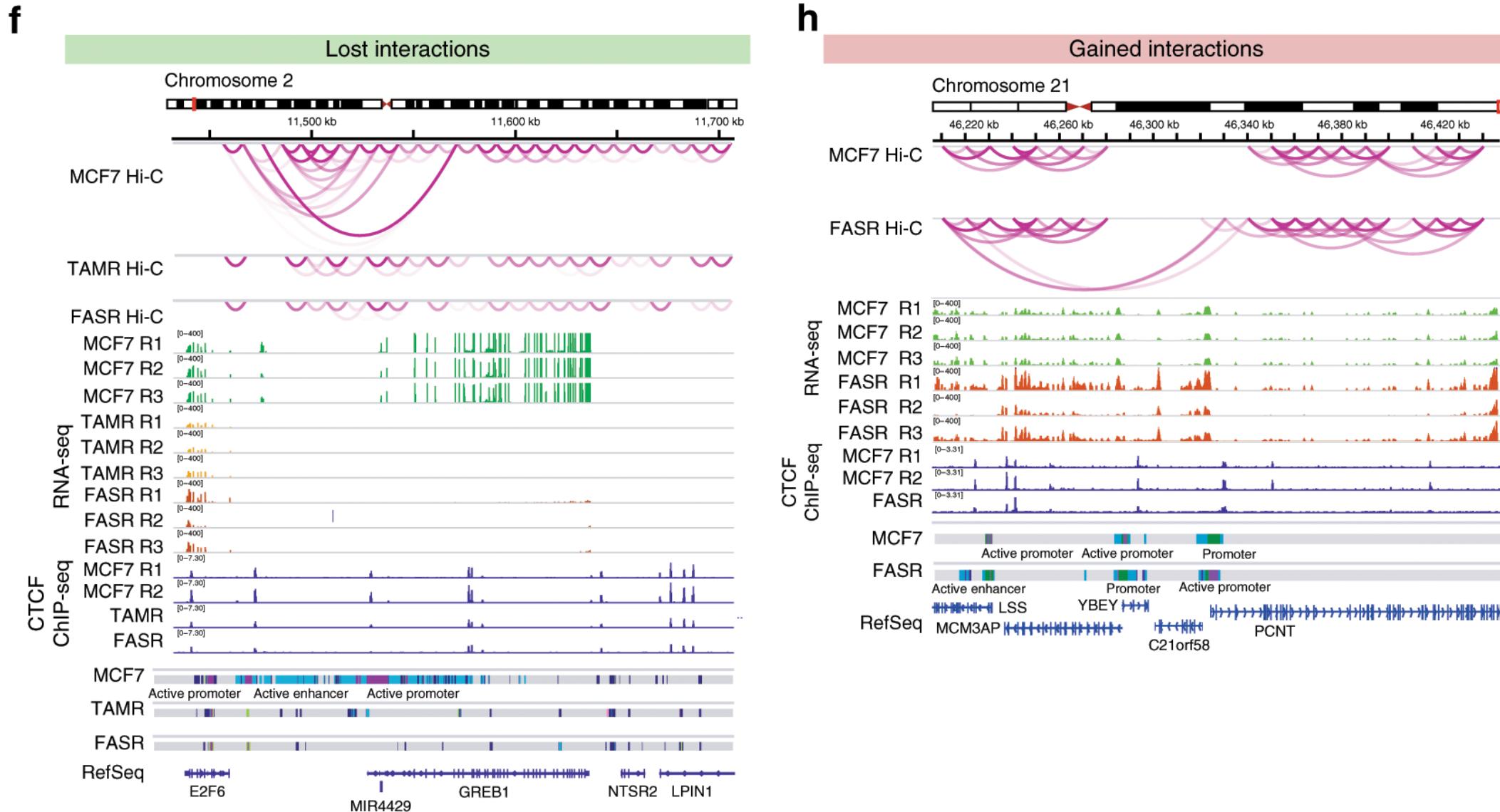
a | Genomic organization of the mouse sonic hedgehog (Shh) locus. In developing limb buds, the ‘zone of polarizing activity regulatory sequence’ (ZRS) enhancer controls the expression of Shh located ~850 kb away.

b | Mice lacking the ZRS Shh limb bud enhancer are born with severely truncated limbs.

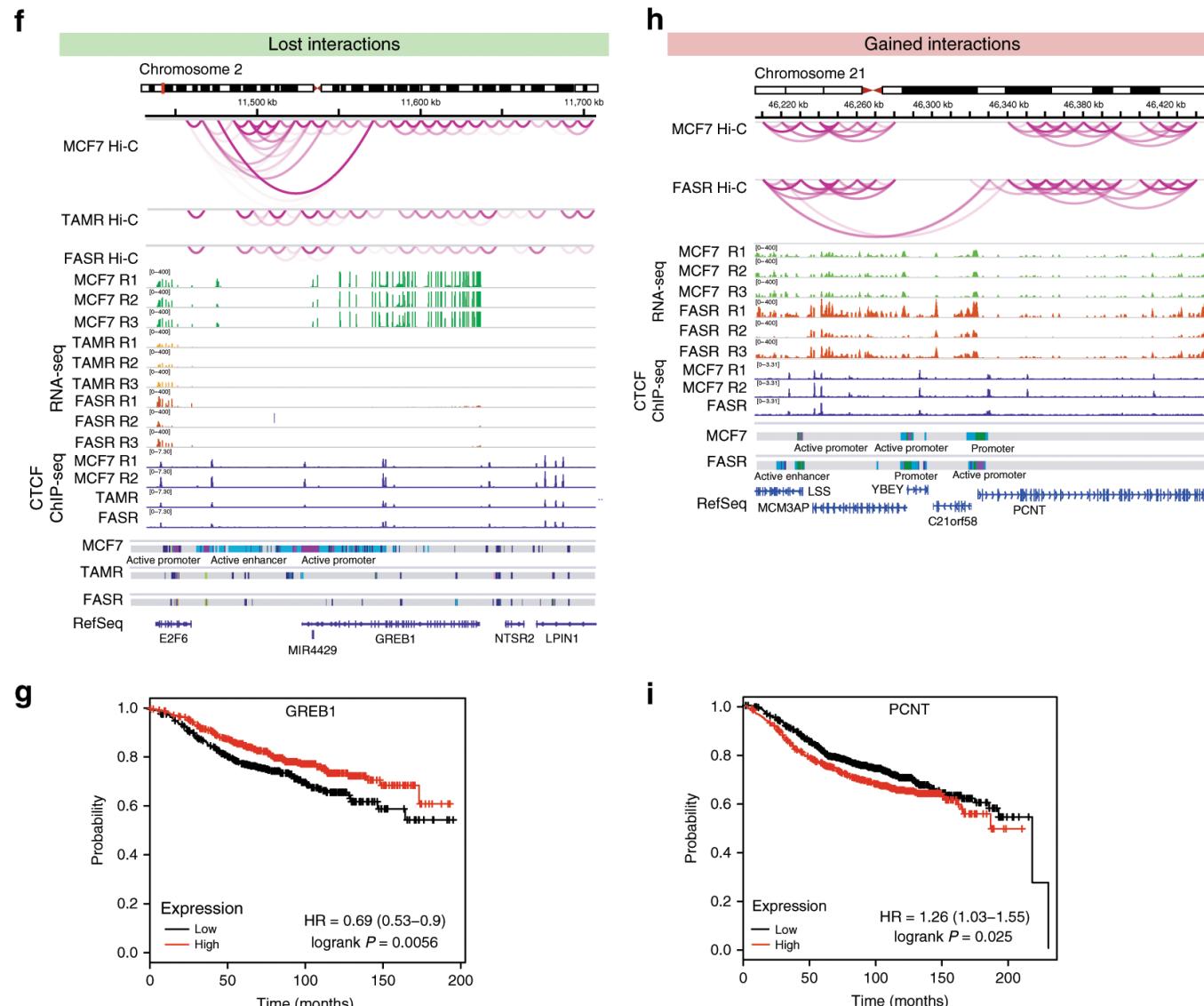
Differential enhancer–promoter interactions and gene deregulation



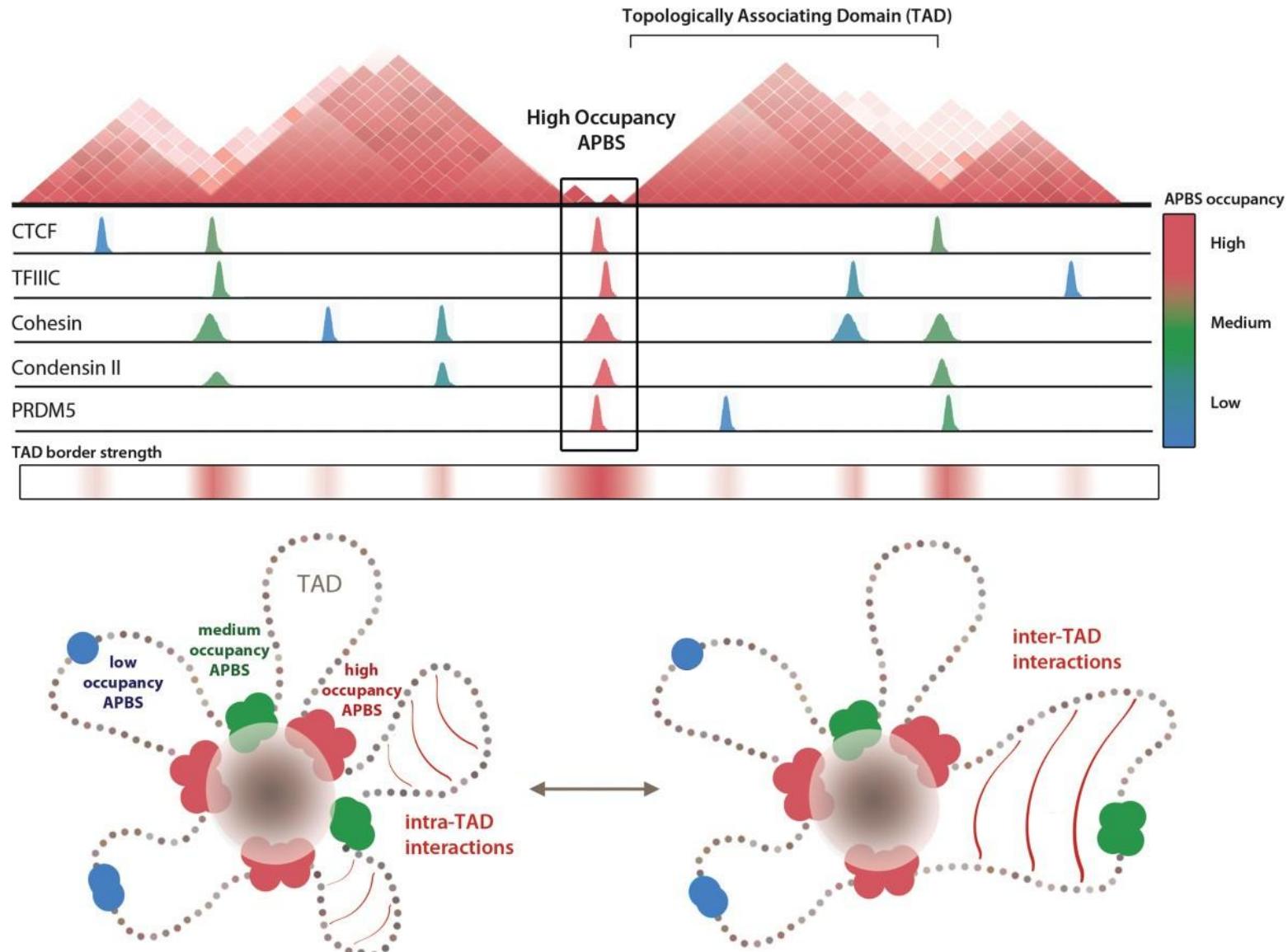
Differential enhancer–promoter interactions and gene deregulation



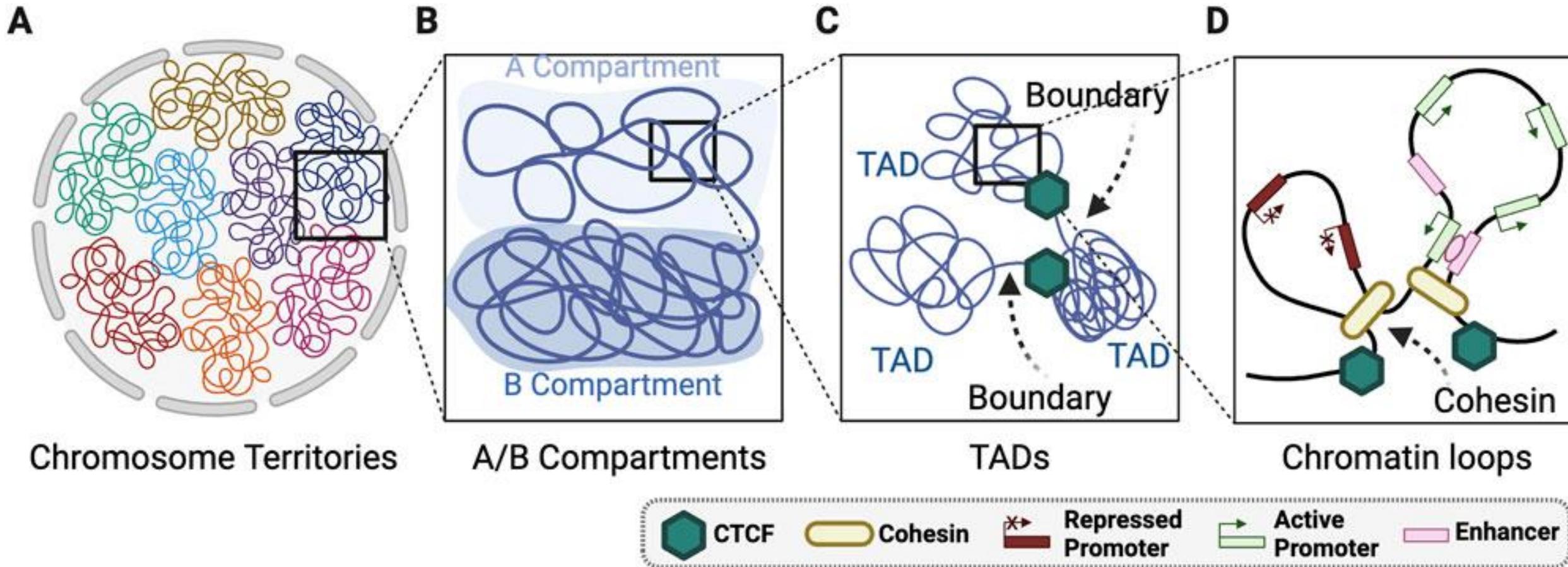
Differential enhancer–promoter interactions and gene deregulation



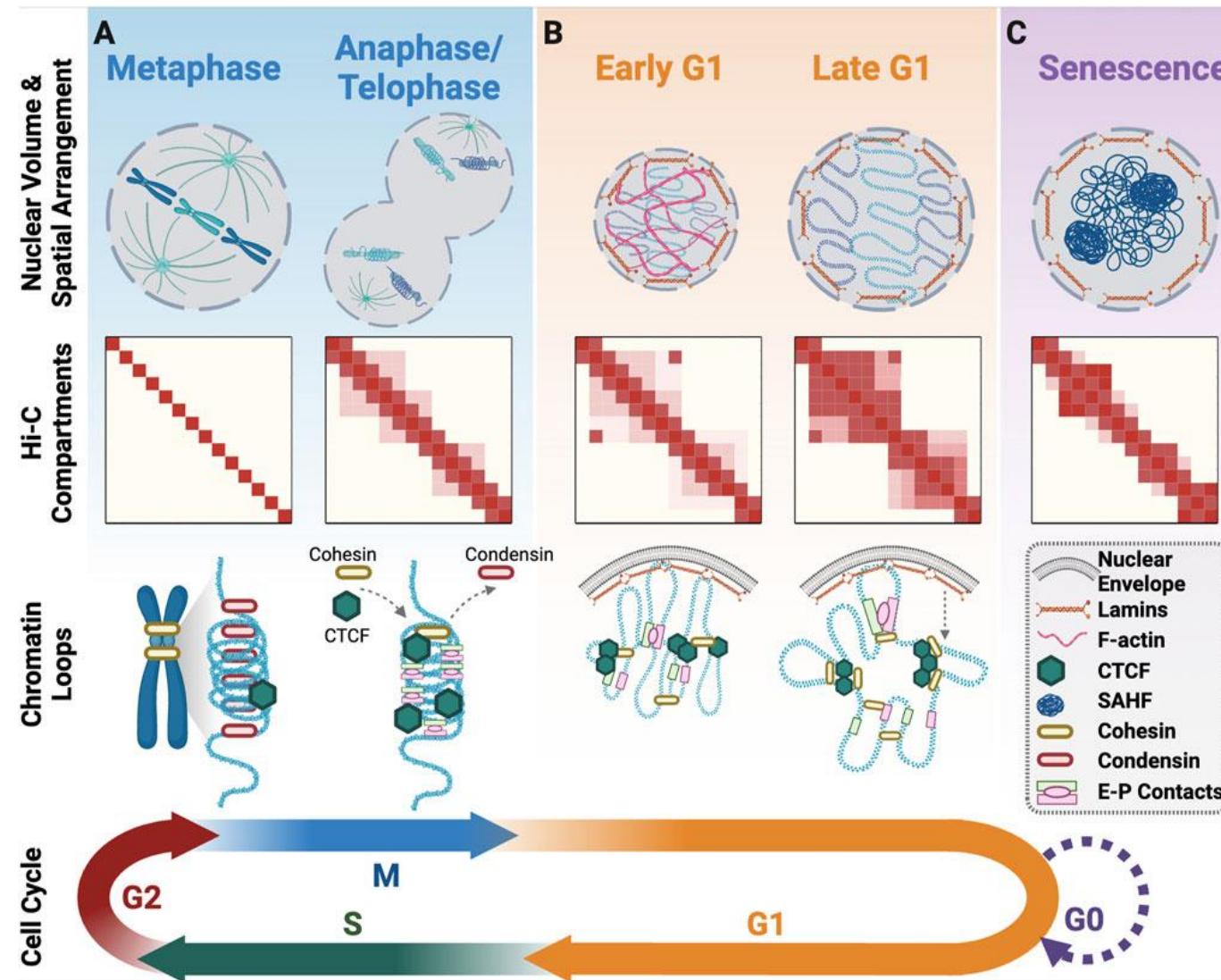
Combinatorial binding of architectural proteins shapes topological domain structure



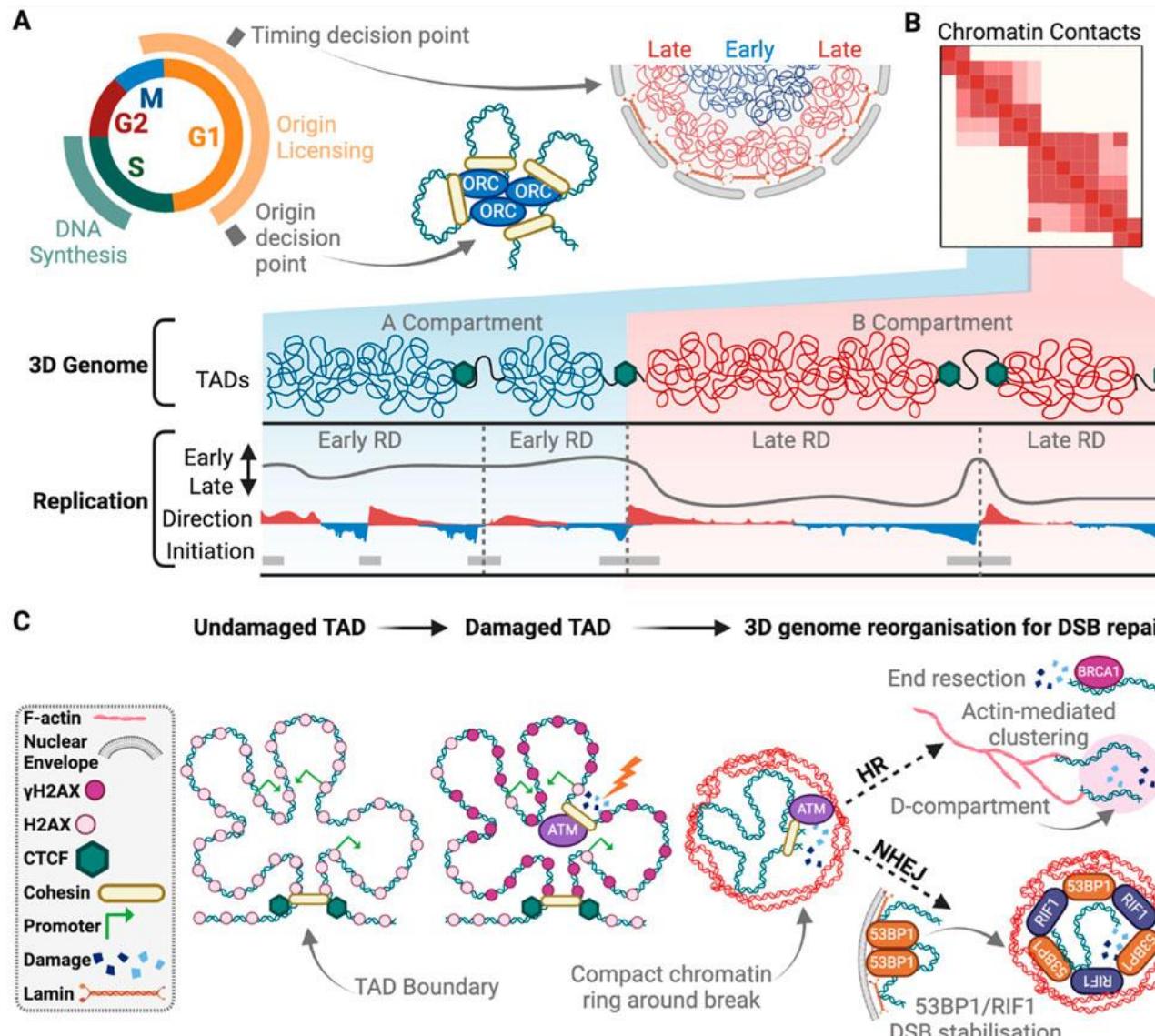
Layers of 3D genome organisation



3D genome organisation upon cell cycle exit and re-entry



The 3D genome in DNA replication and repair



ChIP-seq for Histone modifications, TFBSs and DNasel Hypersensitivity

ChIP-seq overview

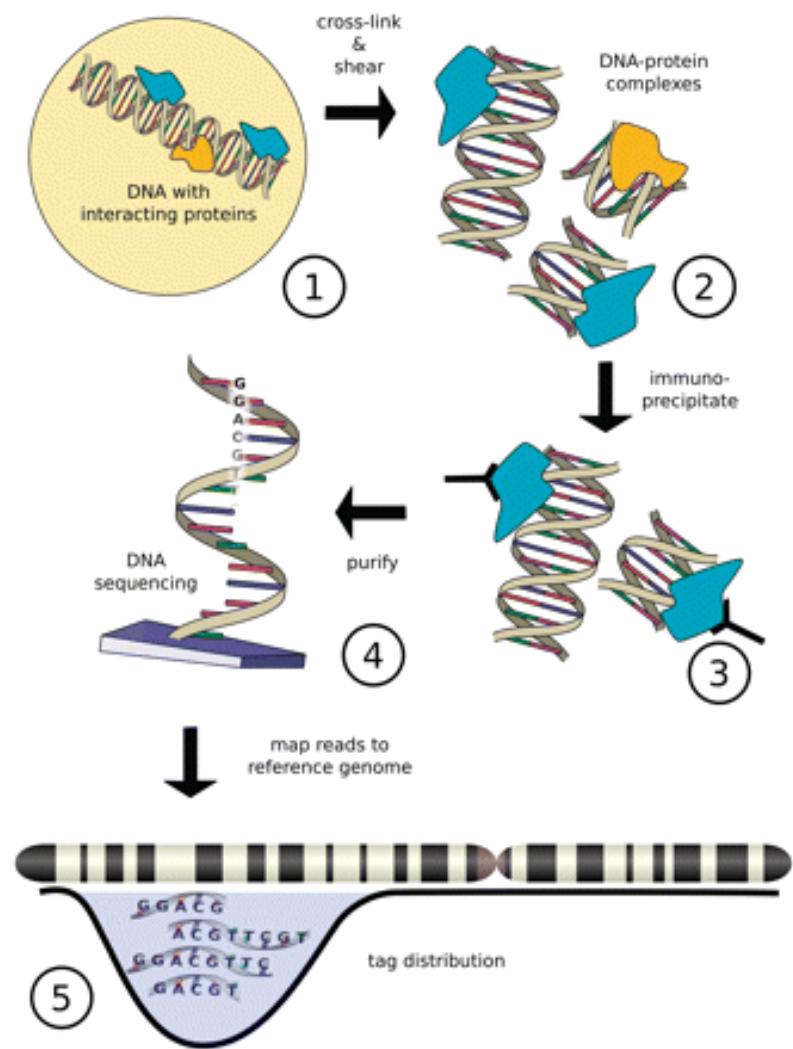
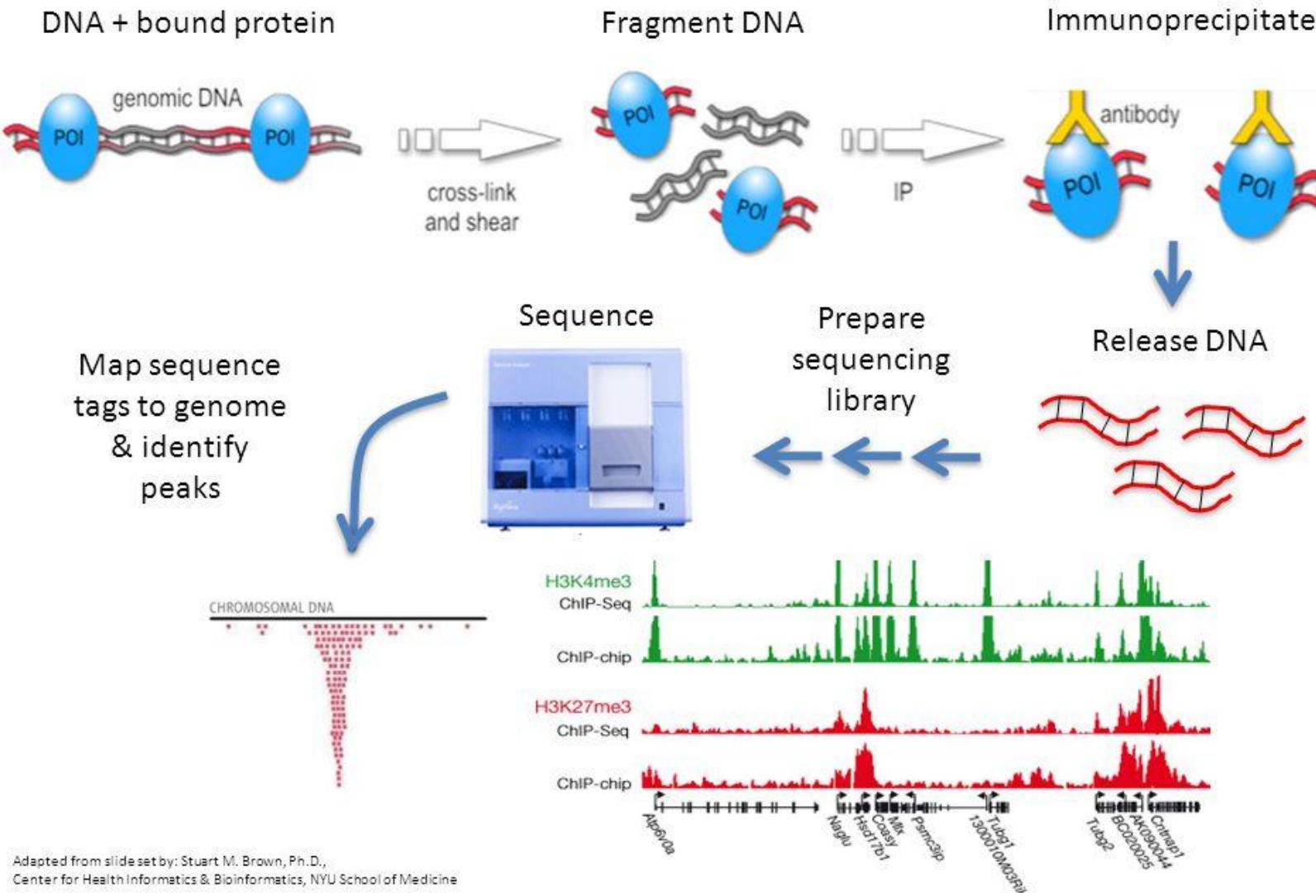
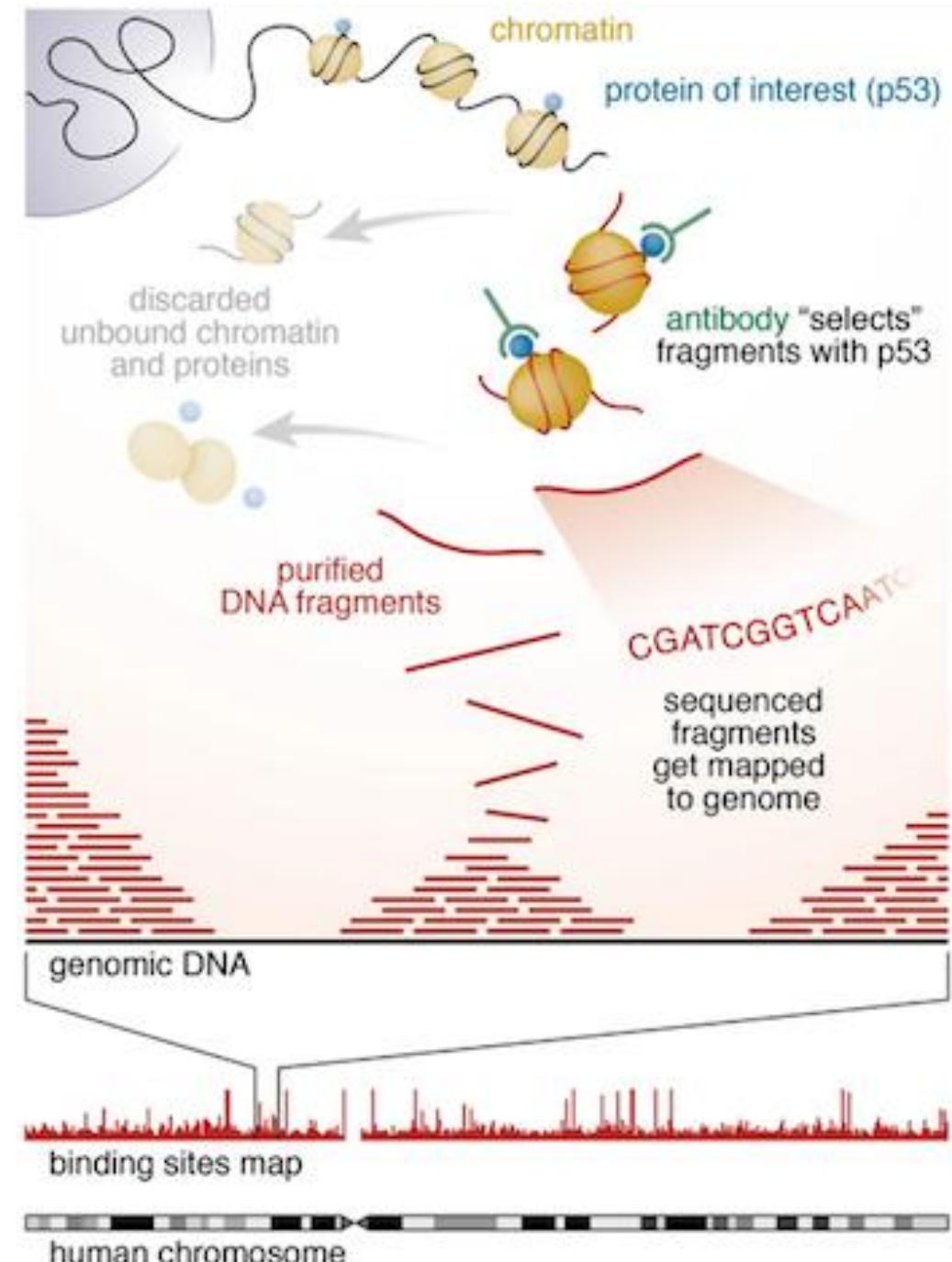
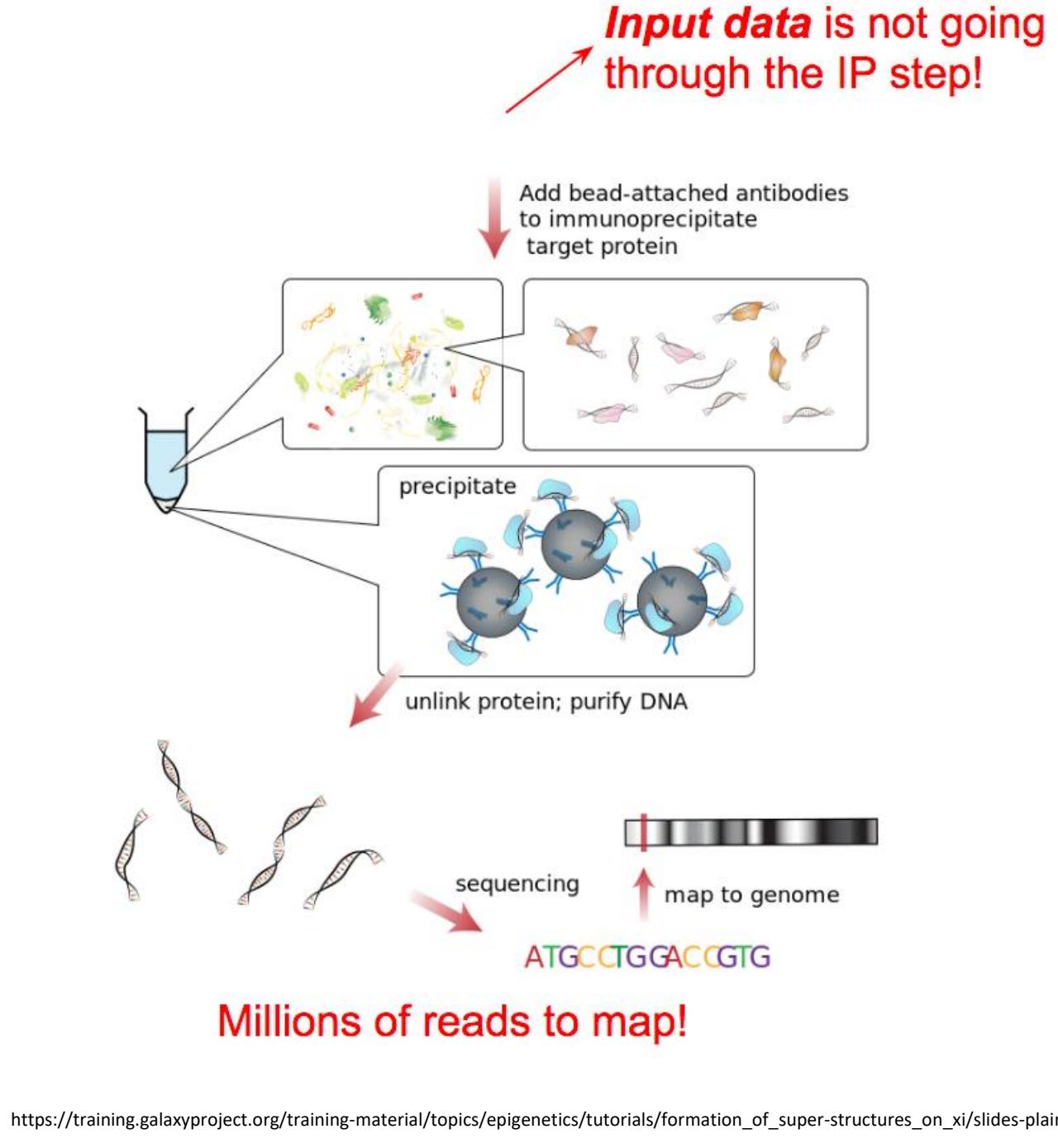
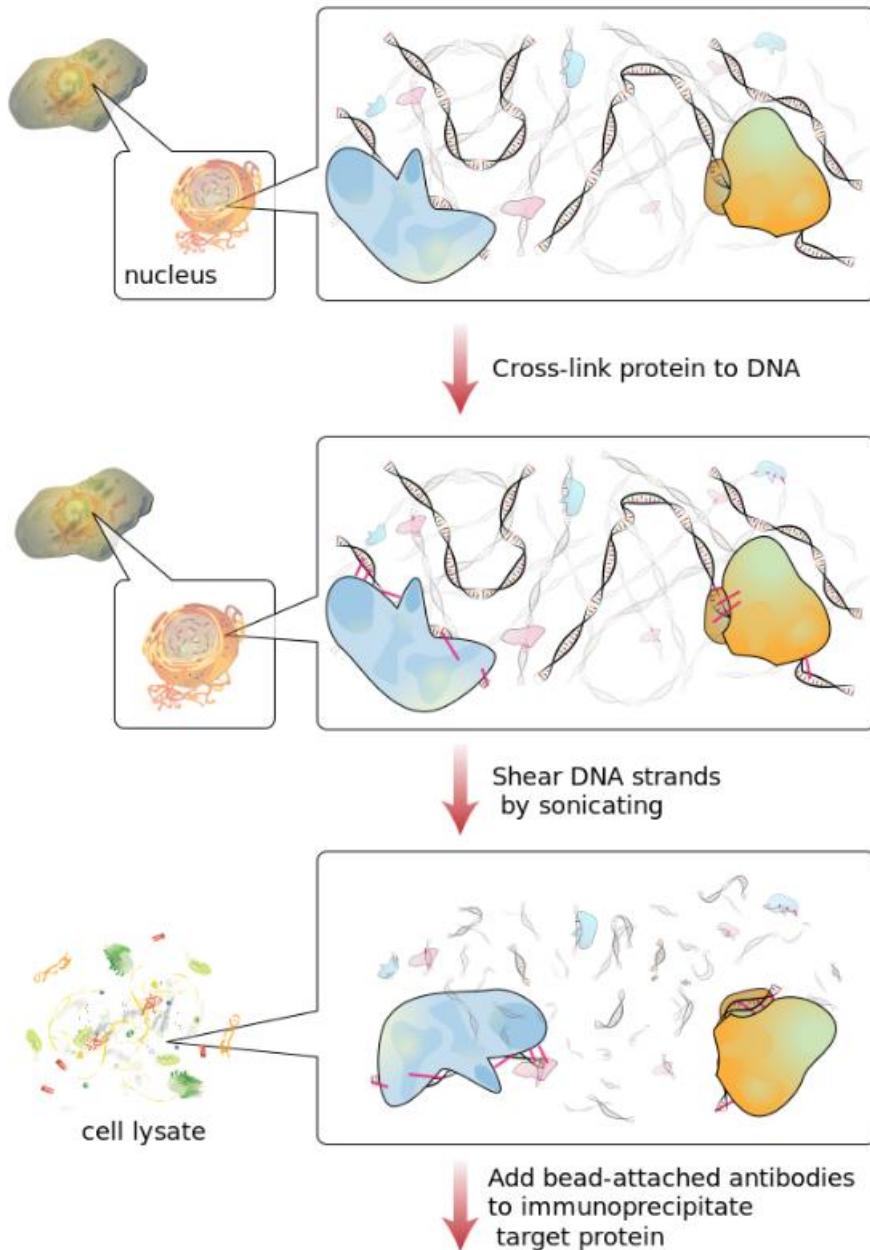


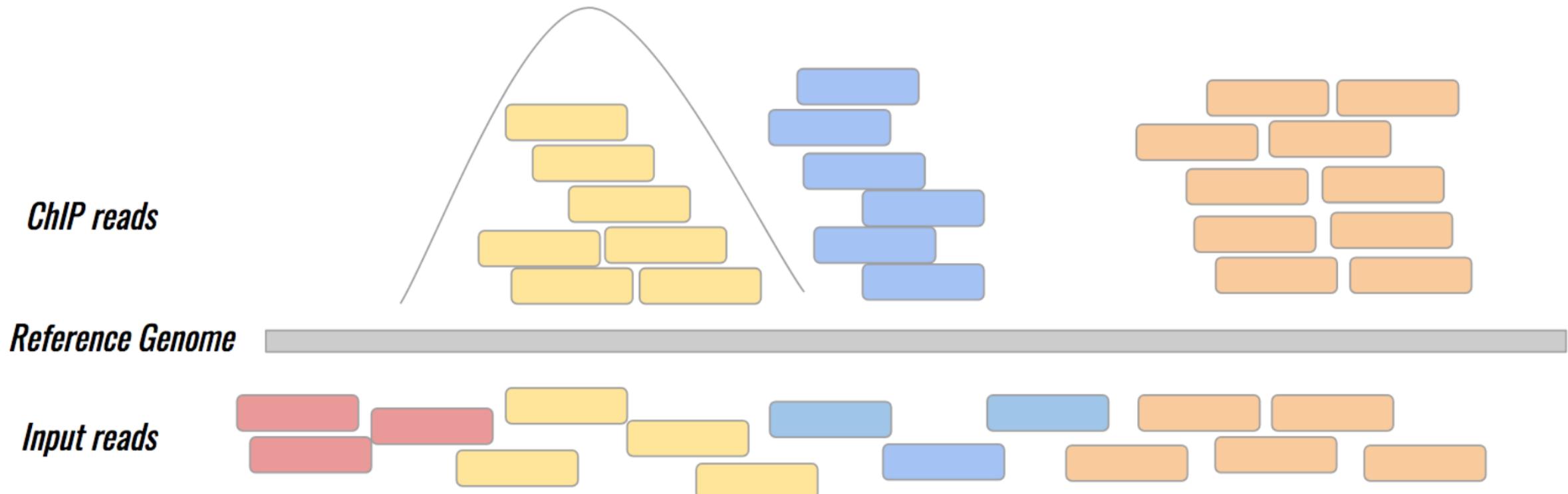
Figure from Szalkowski, A.M, and Schmid, C.D.(2010). Rapid innovation in ChIP-seq peak-calling algorithms is outdistancing benchmarking efforts. *Briefings in Bioinformatics*.

ChIP seq

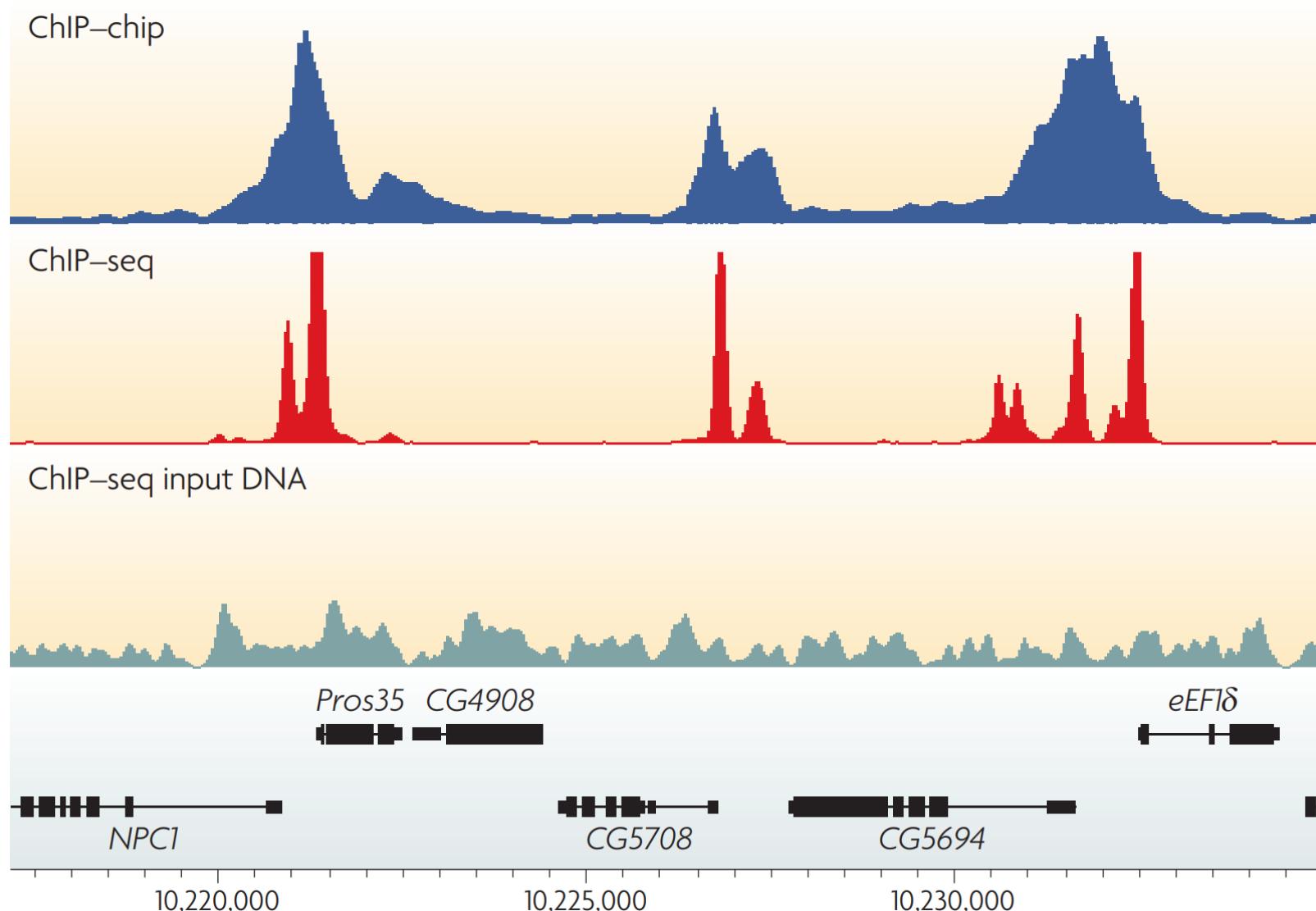




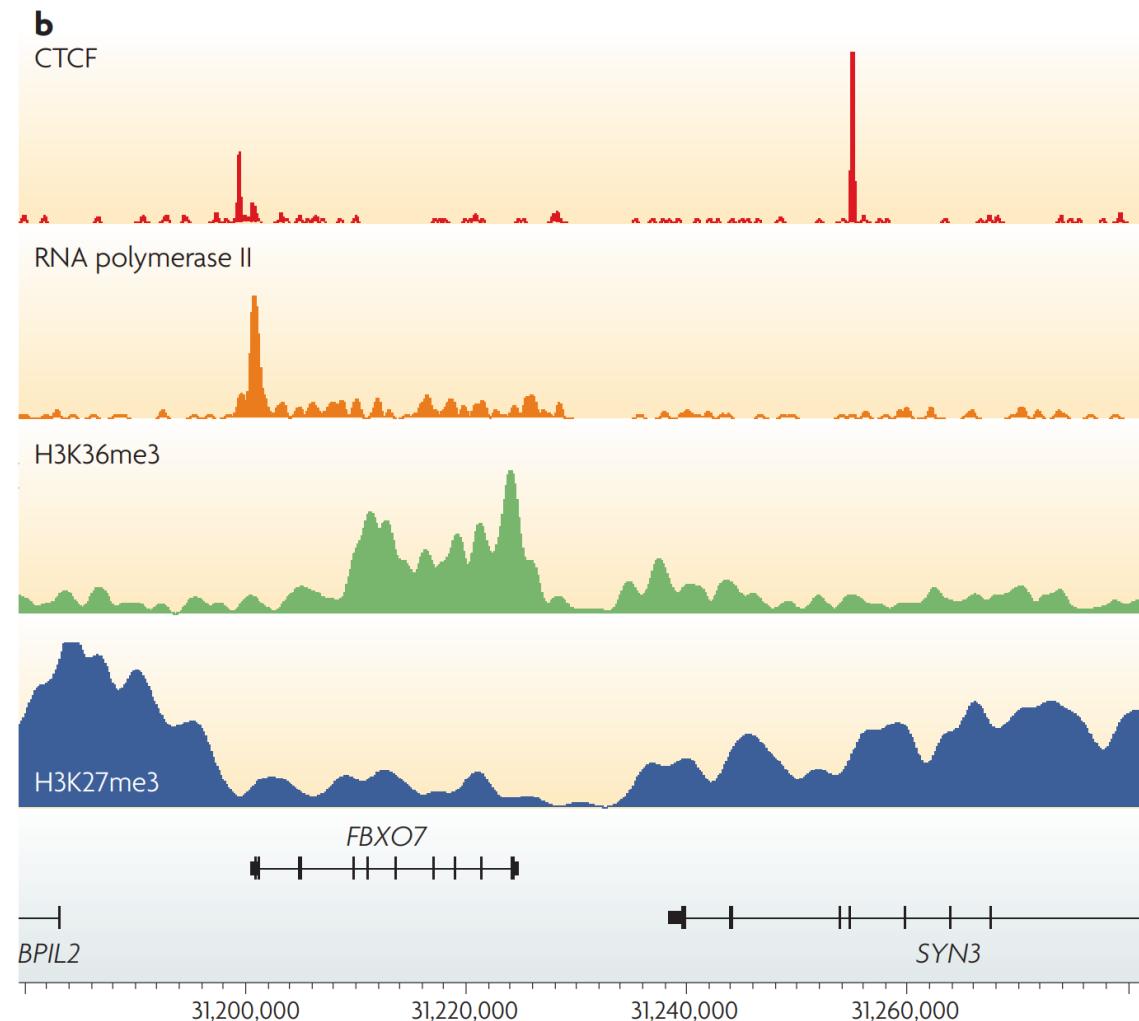
ChIP vs Input



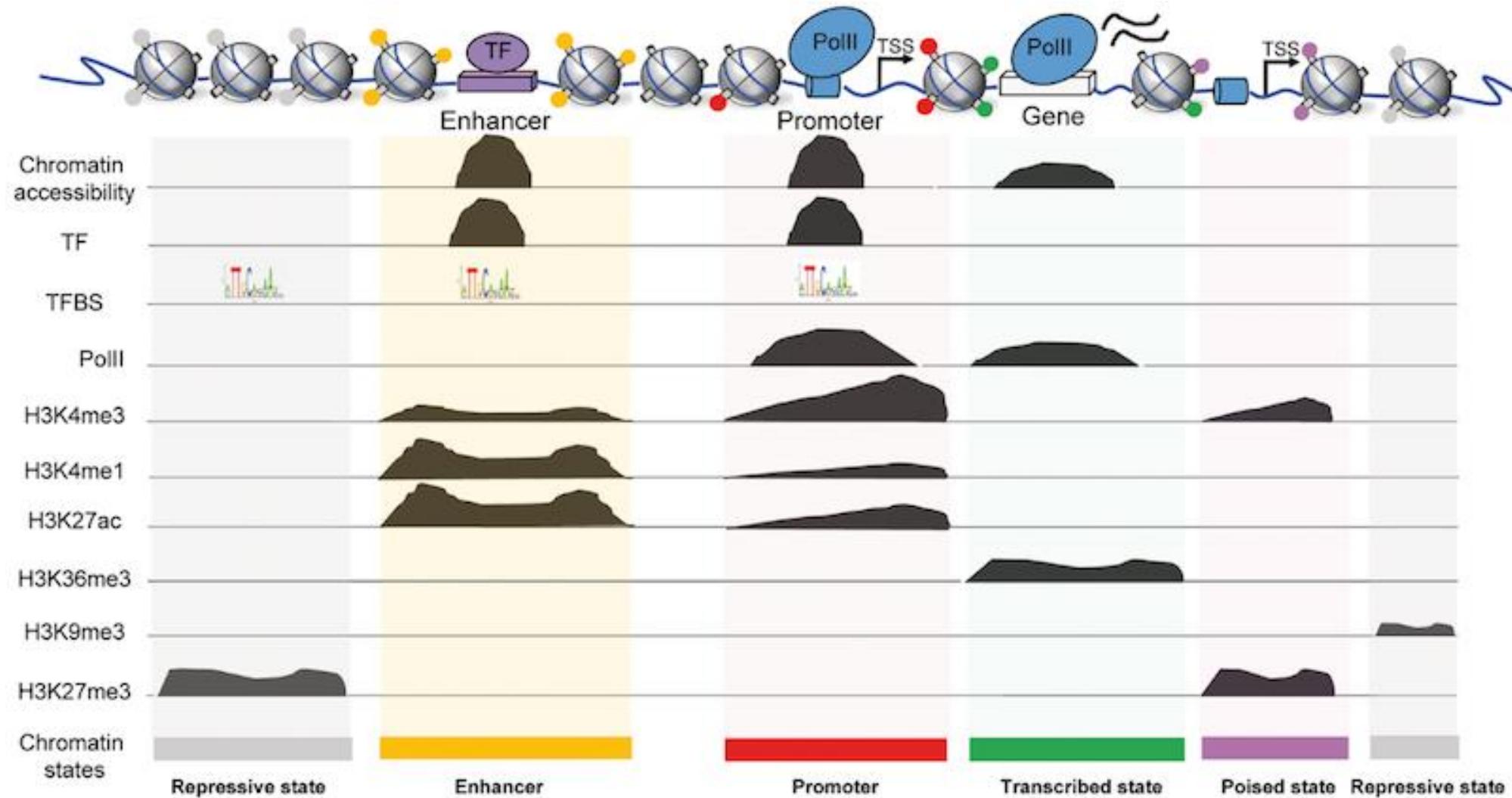
ChIP profiles: Examples of the profiles generated by chromatin immunoprecipitation



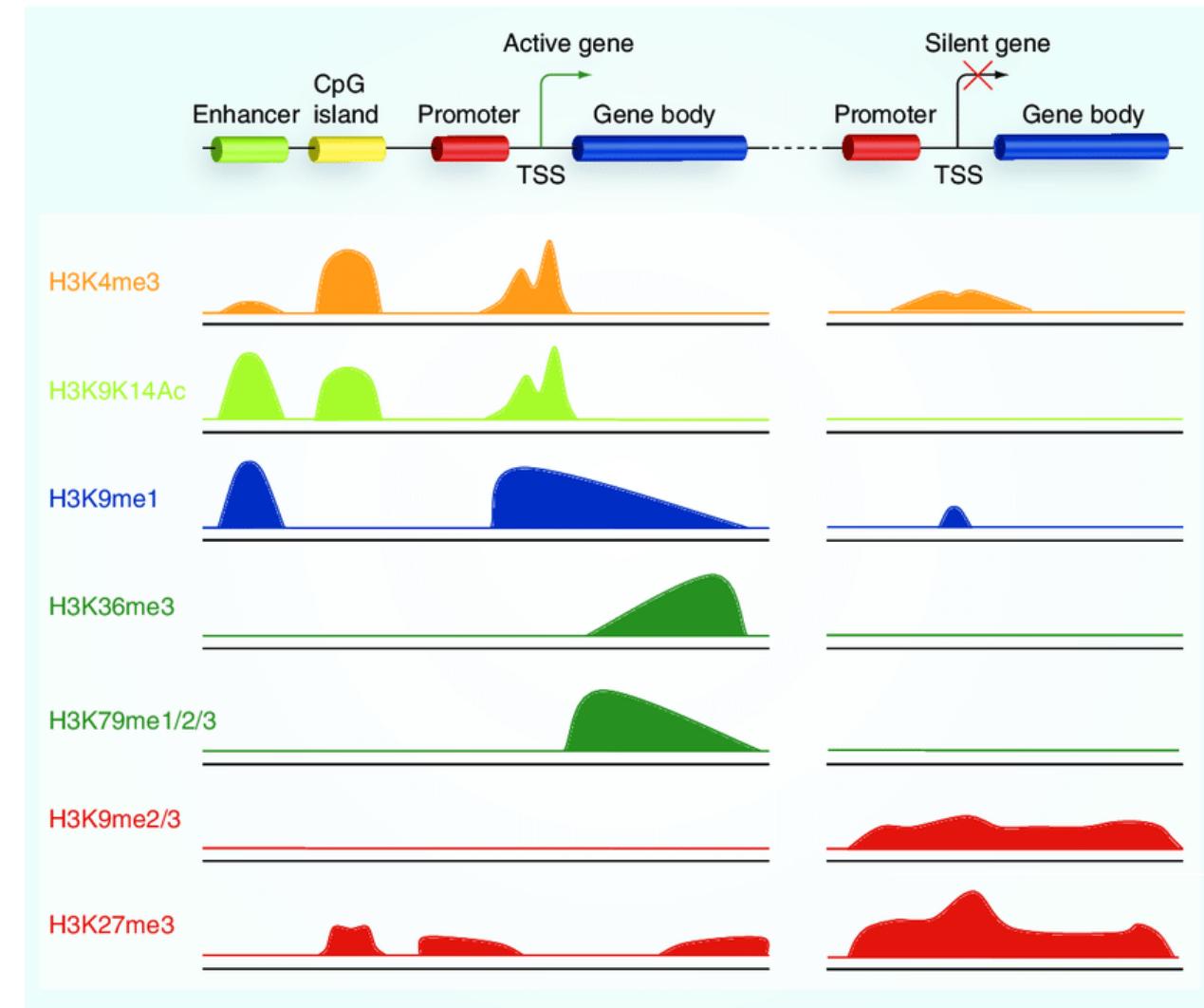
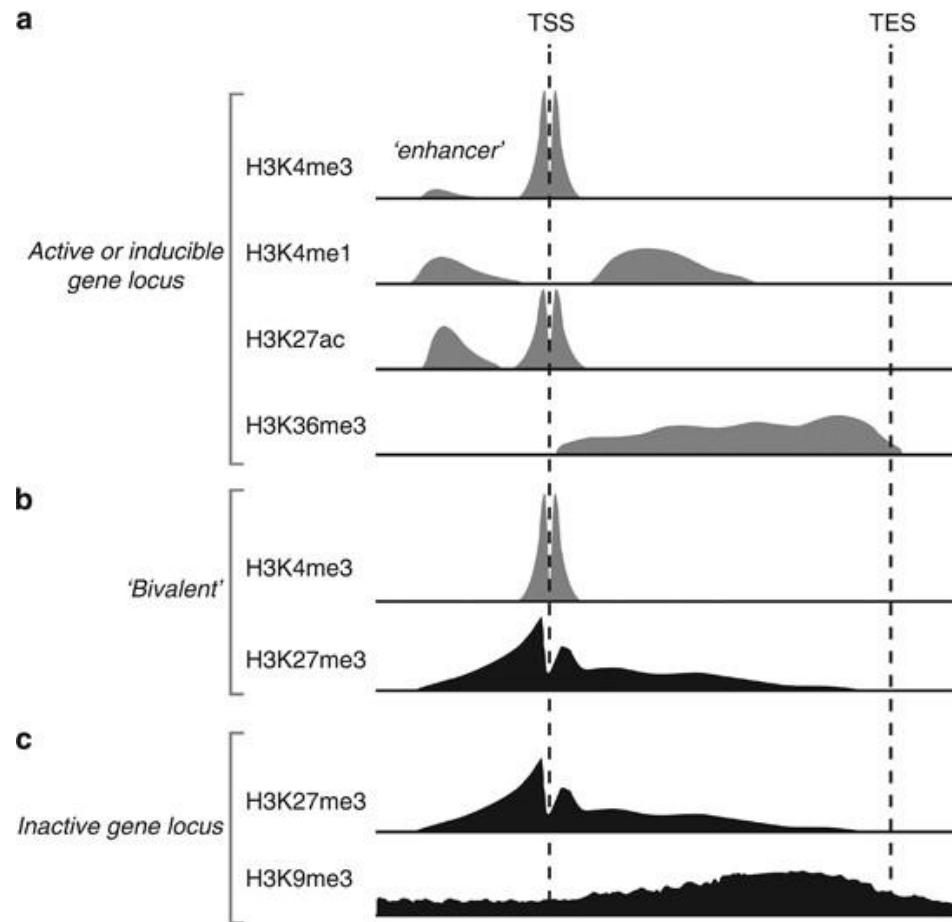
ChIP profiles: Examples of different types of ChIP-seq tag density profiles in human T cells



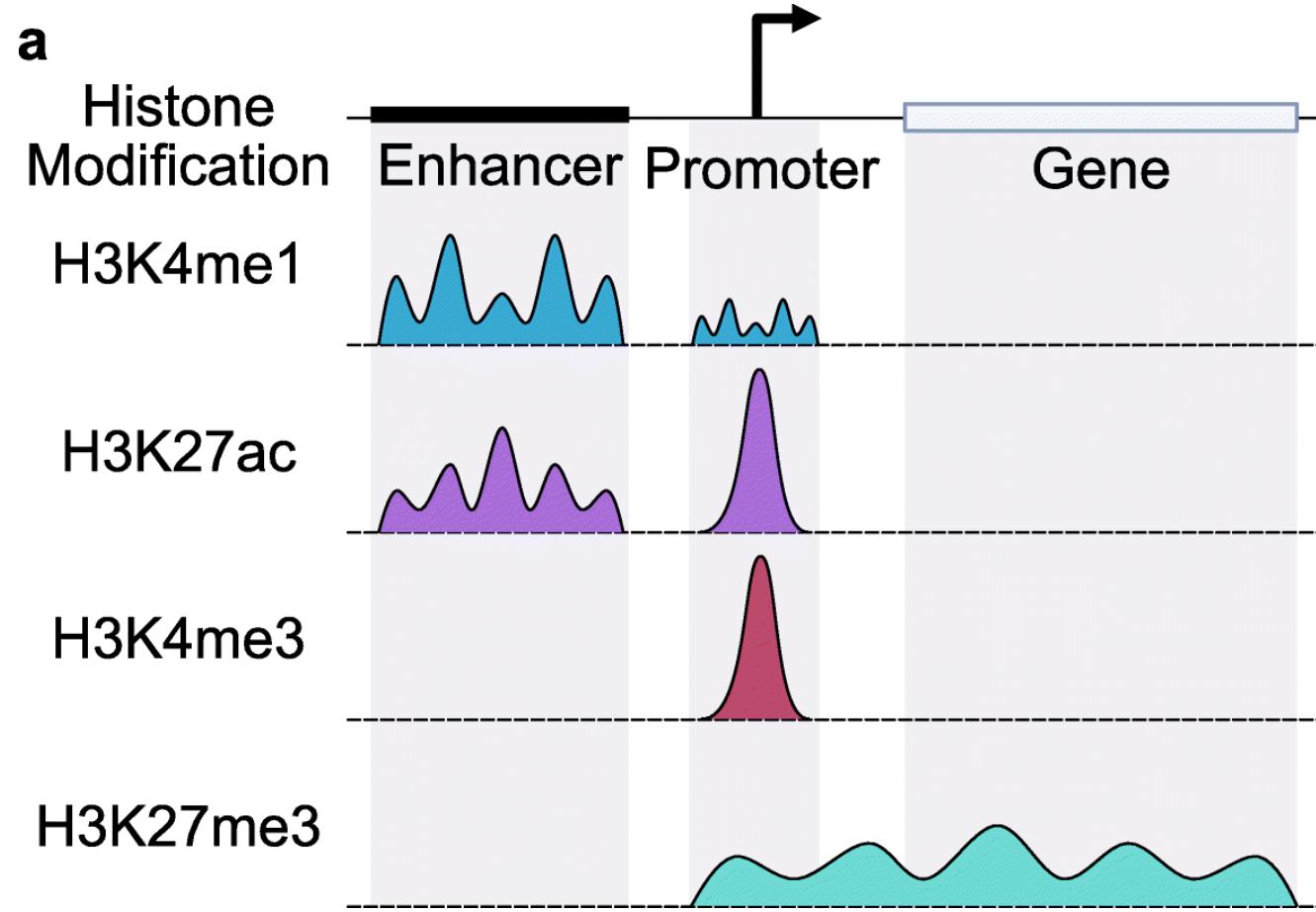
Types of ChIP-seq



Peak types and locations of histone methylation



Peak types and locations of histone modifications

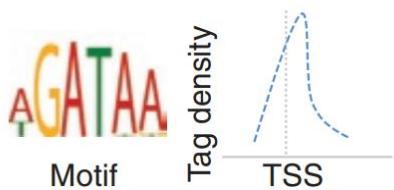




Reference genome



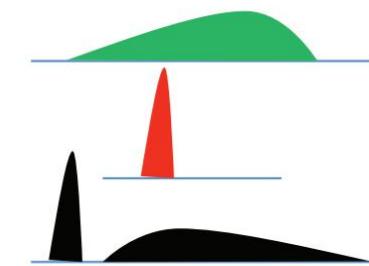
Genome browser



Motif

Tag density

TSS



Sequence alignment

(no longer a bottleneck for data analysis)

Inspect data quality

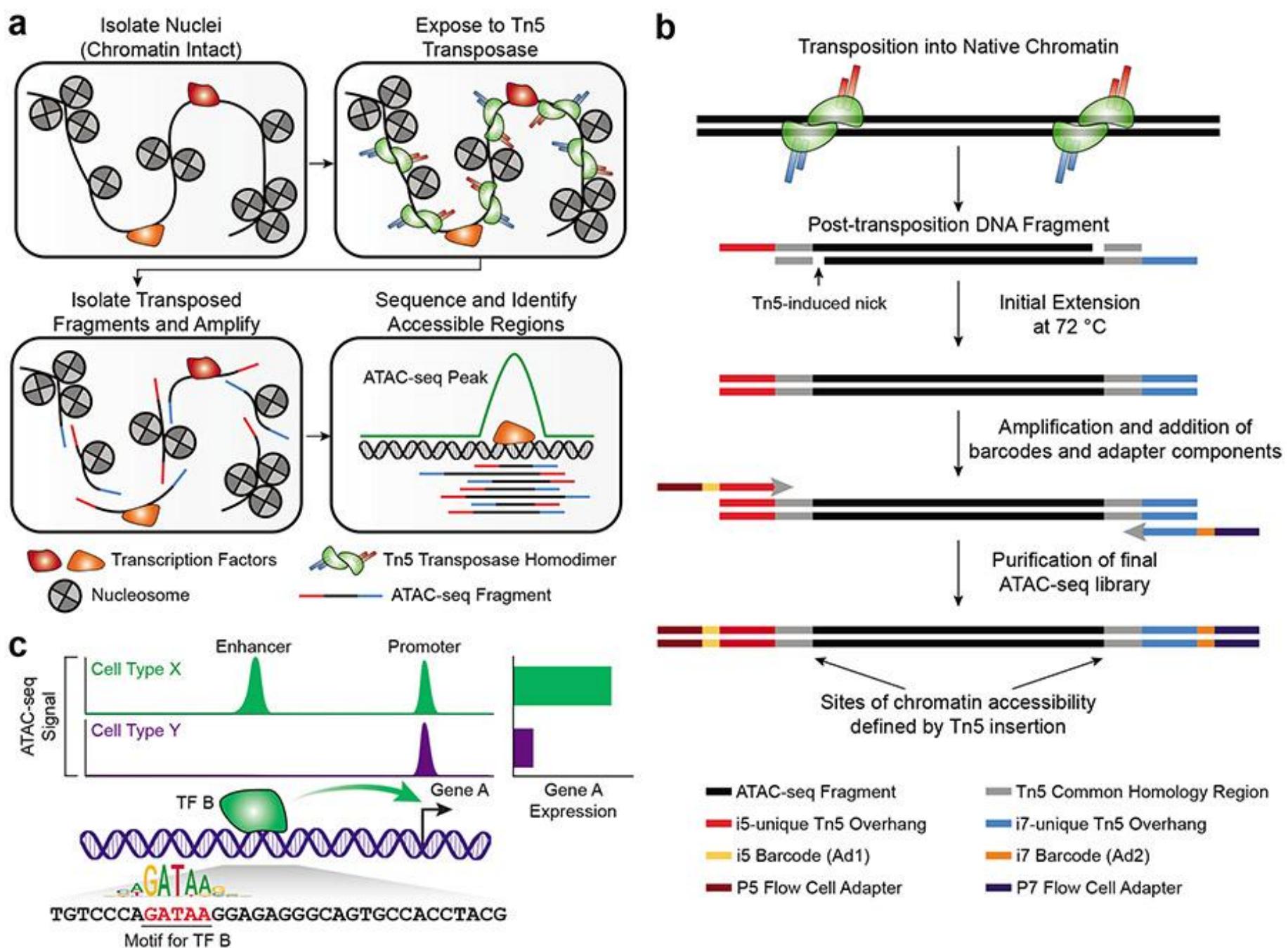
- Check report from vendor-supplied analytical pipeline
- Visually inspect with a genome browser
- Identify motif in tag-enriched region
- Examine the profile at certain genomic features
- Confirm by quantitative PCR

Peak calling (choose the right tool)

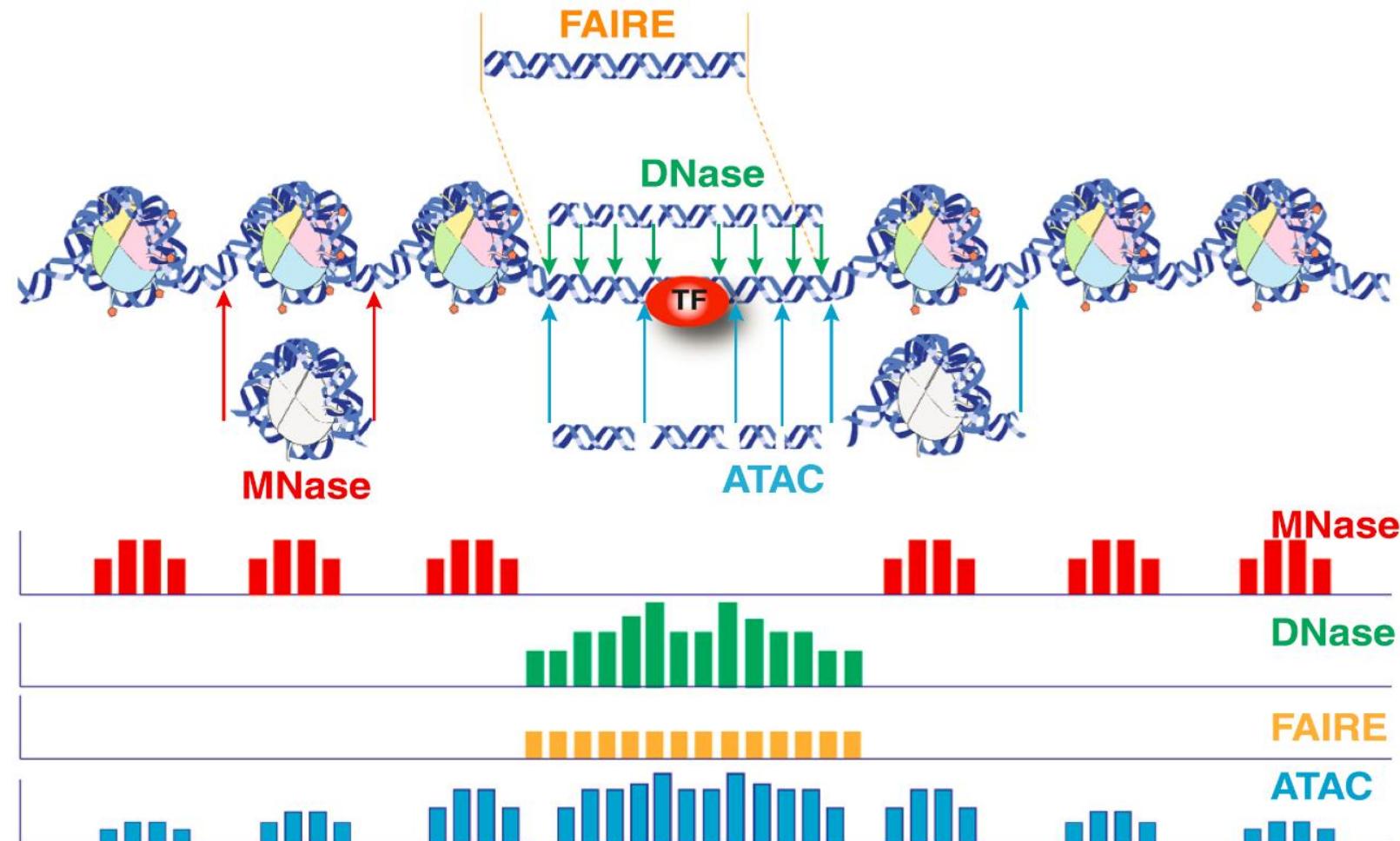
Type of peak	Example	Representative tools
Broad	H3K27me3	CCAT, SICER
Sharp	CTCF	MACS
Sharp & broad	Pol II	ZINBA

Figure 2 Common procedures for ChIP-seq data analysis. After base-calling, short-read sequences are aligned to a reference genome. Data quality is assessed by a combination of various strategies, such as visual inspection with a genome browser, motif identification and confirmation by ChIP and quantitative PCR. The initial inspection or prior knowledge provides information about whether the peaks are broad or sharp or both. Various algorithms (bottom right) have been developed for the identification of peaks of these three groups.

ATAC-seq



ATAC-seq



Population-scale measurements of chromatin accessibility reflect the average accessibility of a heterogeneous collection of single molecules

