DROSOPHILA GENOME (SEQUENCED IN 2000)

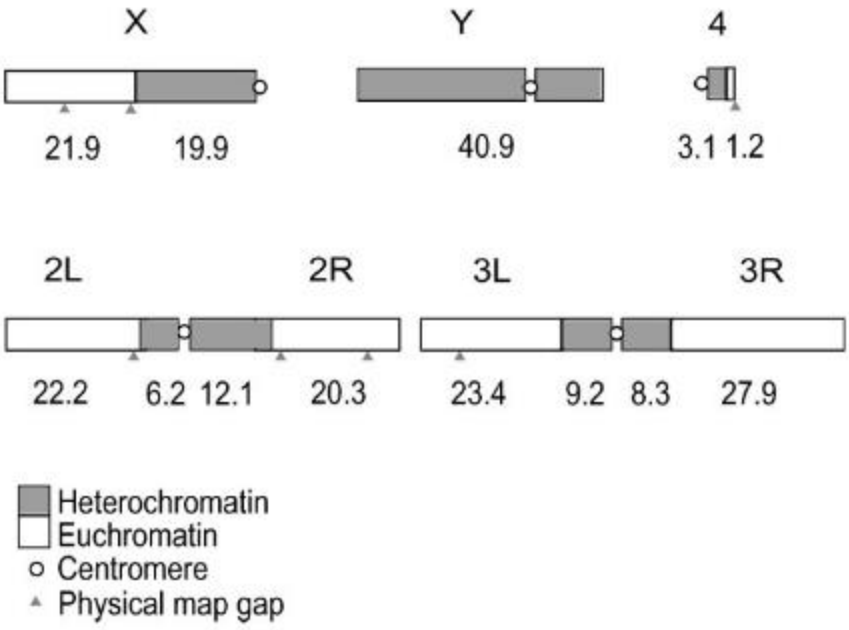
Each of the large chromosomes has a DNA molecule of _5 cm, it has to fit into a nucleus of an average diameter of _5 mm. Therefore, chromosomes must be condensed thousands of times on the linear scale to fit into the nucleus.

At the periphery of the nucleus, chromatin is associated to the nuclear lamina and appears very compact and generally transcriptionally inactive, while closer to the centre of the nucleus chromatin is less compact and enriched in transcriptionally active regions. *Drosophila* chromosomes adopt a polarised orientation, known as the **Rabl configuration,** with centromeres and telomeres clustering separately at two opposite nuclear poles homologous chromosomes are paired throughout the cell cycle.

The nucleolus serves as an anchor for chromosomes, and is surrounded by peri-centromeric heterochromatin. The distance of pericentromeric heterochromatin regions from the periphery varies by chromosome, with chromosomes 4 and X heterochomatin more peripheral relative to Pericentromeric regions of other chromosomes.

- *Drosophila* genome is relatively small 180 Mb with only three pairs of autosomes, a third of which is centric heterochromatin (about 60 Mb) and one pair of sex chromosome. The euchromatin contains about 98% of the protein-coding genes in the genome; the heterochromatin is largely composed of simple sequence repeats.
- The heterochromatin consists mainly of short, simple sequence elements repeated for many megabases, occasionally interrupted by inserted transposable elements, and two large blocks of ribosomal RNA genes.
- It is known that there are small islands of unique sequence embedded within heterochromatin.
- Within heterochromatin about 21% of this DNA consists of highly repetitive satellite DNA that is centered in centromere and Y chromosome.

- In *Drosophila*, the genome contains roughly 13,600 genes and is organized in territories, **A/B compartments** and **topologically associating domains (TADs).**
- 'TADs',(length 100 kb on average) are defined as chromosomal regions displaying a high self-interaction frequency delimited by a dramatic decrease of interaction frequencies with the regions on either side of the domains that defines their boundaries.
- Genome analysis also revealed a segregation of the genome into two types of compartments, called **A** and **B compartments**.
- The A compartment contains gene-rich and transcriptionally active open chromatin, while the B compartment is gene-poor and enriched for closed chromatin regions. Compartments tend to interact more frequently within their own class. Both comportments are large multi mega base-sized regions.
- About 50 families of transposable elements occupying about 9% of DNA is present in drosophila genome. They are roughly 2-9 kb in length and their number of copies varies from 10 to100 among different strains. These are mostly found at centromeric and telomeric regions along within euchromatic part in a scattered fashion.
- The 120 Mb of euchromatin is on two large autosomes and the X chromosome; the small fourth chromosome contains only ∼1 Mb of eu-chromatin.



❖ A FINER APPROACH

Recently, genome-wide chromatin contacts have been determined for16-18 hr Drosophila embryos using the Hi-C technique. The euchromatin genome (chromosomes 2, 3, 4, and X) was divided into 1169 physical domains based on Hi-C interaction profiles. These physical domains were assigned to four functional classes based on their average epigenetic signatures: **Null, Active**, **Polycomb-Group (Pc-G)** and **HP1/Centromere**.

Inactive TADs roughly correspond to regions of strong attachment to the nuclear lamina **[lamina associated domains (LADs)],** whereas active TADs are characterized by lower frequencies of lamina association.

Furthermore, TADs correlate even better with domains of a defined timing of DNA replication during the S phase of the cell cycle; with active TADs equivalent to early replicating domains and heterochromatic TADs equivalent to late replicating domains.

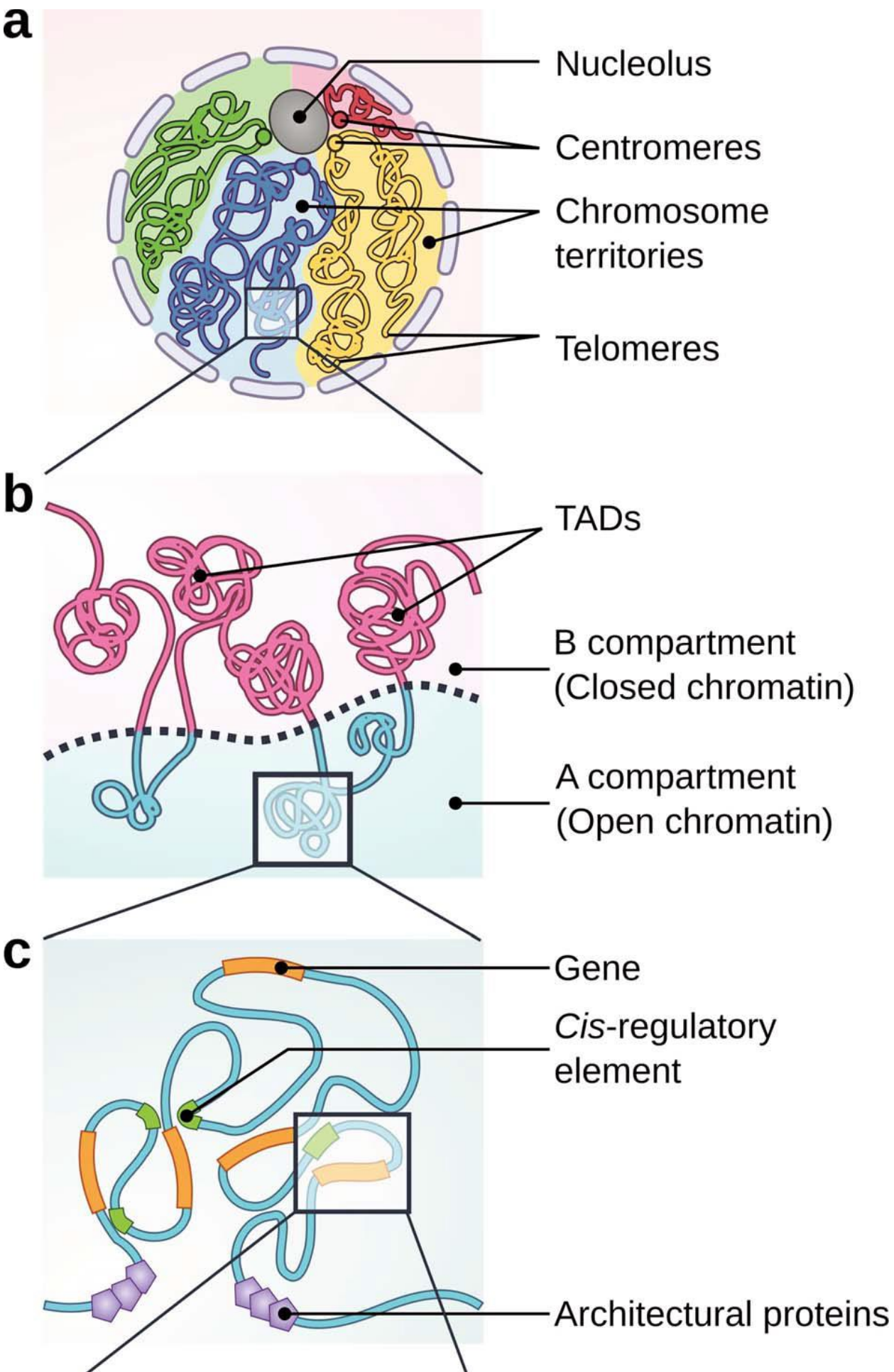
Three types of inactive TADs have been identified. The first corresponds to Polycomb repressed loci enriched in histone H3 tri methylated at lysine27(H3K27me3). Polycomb TADs represent 10% of the fly genome and

contain a large number of developmental genes, many of which encode transcription factors involved in patterning.

A second type of silent TADs contains heterochromatin. This is mainly located at peri-centromeric regions as well as telomeric regions of the chromosomes.

A final type of repressive chromatin domain is defined as void or **black chromatin**. This contains up to 50% of euchromatin of the genome, characterized by low or no transcriptional activity and low or absent histone modifications of any kind. It seems that black chromatin may represent a passive inactive state which, upon selective recruitment of specific components depending on developmental cues, can switch into a Polycomb, heterochromatic, or an active state.

Conversely, the relatively decondensed active domains might correspond to low nucleosome interactions. Experimentally, the acetylation of histones is known to promote chromatin unfolding by decreasing nucleosome–nucleosome interactions, as the neutralisation of positive charges on histone tails on a nucleosome hinders interaction with the acidic patch of another nucleosome


a
Nucleolus
Centromeres
Chromosome territories
Telomeres
b
TADs
B compartment (Closed chromatin)
A compartment (Open chromatin)
c
Gene
Cis-regulatory element
Architectural proteins