

The Structure and Function of the Eukaryotic Chromosome

Polytene Chromosome Structure and Sequence Organization

A eukaryotic chromosome is a single, enormously long DNA molecule packaged with proteins into a compact, functional structure. Understanding this packaging - and the specialised, naturally "unpacked" polytene chromosomes found in some insect tissues - along with how the underlying DNA sequence itself is organised (unique versus repetitive), together explain both how chromosomes are built and how they are read out during gene expression.

1. Structure and Function of the Eukaryotic Chromosome

A single human chromosome can contain several centimetres of DNA that must fit into a nucleus only a few micrometres across. This is achieved through successive levels of packaging, each adding further compaction while - critically - remaining reversible, so the cell can still access its DNA for transcription, replication and repair.

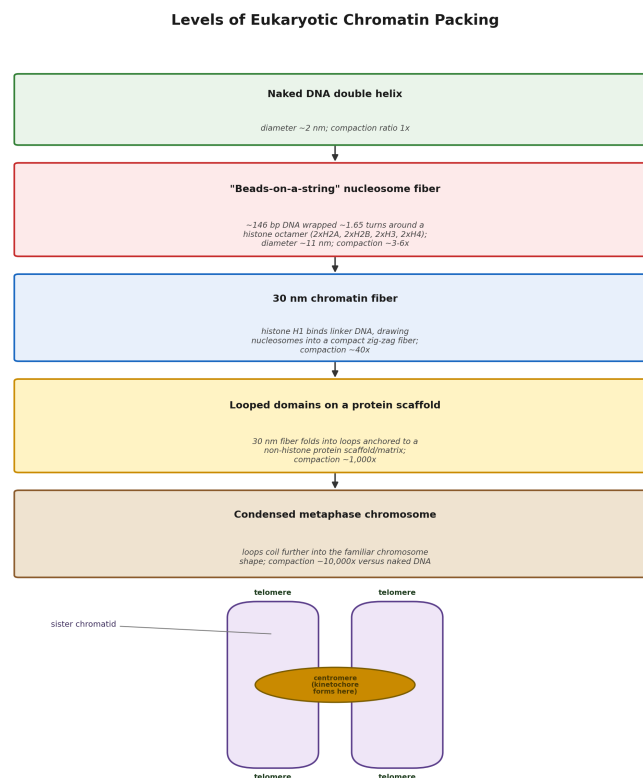


Fig. 1 - From the naked DNA double helix to a condensed metaphase chromosome: successive levels of chromatin packing, and the key functional landmarks of the finished chromosome.

1.1 The nucleosome: the basic packaging unit

The first level of packaging is the **nucleosome**: about 146 base pairs of DNA wrapped in ~1.65 left-handed turns around a disc-shaped **histone octamer** (two molecules each of histones H2A, H2B, H3 and H4). Short stretches of **linker DNA** (roughly 20-80 bp) connect successive nucleosomes, giving the classic "beads-on-a-string" appearance and roughly a 3-6 fold compaction of the DNA. The linker histone **H1** then binds where DNA enters/exits each nucleosome and helps draw the array into a more compact **30 nm chromatin fiber** (~40-fold compaction). This fiber is further organised into **looped domains** anchored to a non-histone protein **scaffold/matrix** (~1,000-fold compaction), which coil further to produce the familiar condensed **metaphase chromosome** - an overall compaction of roughly 10,000-fold relative to naked DNA.

1.2 Histone tail modifications

Core histones have flexible N-terminal "tails" that project from the nucleosome and are subject to reversible chemical modifications - **acetylation**, **methylation** and **phosphorylation** among others. These modifications alter chromatin compaction directly (e.g. acetylation tends to loosen chromatin) and also serve as docking marks that recruit specific regulatory proteins - together forming a combinatorial "histone code" that helps determine whether a region of chromatin is transcriptionally active or silent.

1.3 Functional landmarks of the chromosome

- **Centromere**: a specialised chromosomal region that directs assembly of the **kinetochore**, the protein complex that attaches the chromosome to spindle microtubules so sister chromatids can be pulled apart accurately during cell division.
- **Telomere**: repetitive DNA sequences that cap each chromosome end, allowing complete replication of the chromosome tip and protecting the end from being mistaken for a damaged DNA break requiring repair.
- **Origins of replication**: multiple sites along each chromosome where DNA replication begins, allowing the whole chromosome to be copied within a practical timeframe.

1.4 Euchromatin vs heterochromatin

Chromatin exists in two broad functional states. **Euchromatin** is relatively decondensed, gene-rich and transcriptionally active or potentially active. **Heterochromatin** remains highly condensed throughout the cell cycle, is largely gene-poor and transcriptionally silent, and is concentrated at structural regions such as centromeres and telomeres ("constitutive" heterochromatin) or can form reversibly over specific genes to silence them ("facultative" heterochromatin).

Key points

- Chromatin packaging is not just for compaction - each level (nucleosome positioning, 30 nm fiber state, loop attachment) is itself a regulatory layer that can be locally opened or closed to control access to the DNA.
- The centromere is defined more by its associated kinetochore protein architecture than by a strict DNA sequence in many species, including humans.

2. Polytene Chromosome Structure

Polytene ("many-stranded") chromosomes are giant, naturally banded chromosomes found in certain highly metabolically active, non-dividing larval tissues of dipteran insects - most famously the salivary glands of *Drosophila* - as well as in Malpighian tubules, gut, fat body and ovarian nurse cells. Their size and easily visible banding pattern made them one of the earliest tools for physically mapping genes to chromosome positions.

2.1 Formation by endoreplication

Polytene chromosomes form through **endoreplication (endomitosis)**: the cell's DNA undergoes repeated rounds of the S phase (in *Drosophila*, roughly 9-10 rounds) without any intervening mitosis or cytokinesis - no spindle forms and the nuclear envelope stays intact throughout. Rather than separating, the replicated sister chromatids - and, unusually, the homologous maternal and paternal chromosomes too - remain tightly aligned side by side in exact register, a phenomenon called **somatic pairing**. The result is a single giant chromosome bundling roughly a thousand identical DNA strands, which is why *Drosophila* salivary gland nuclei show only 4 polytene chromosomes instead of the diploid complement of 8 chromosomes.

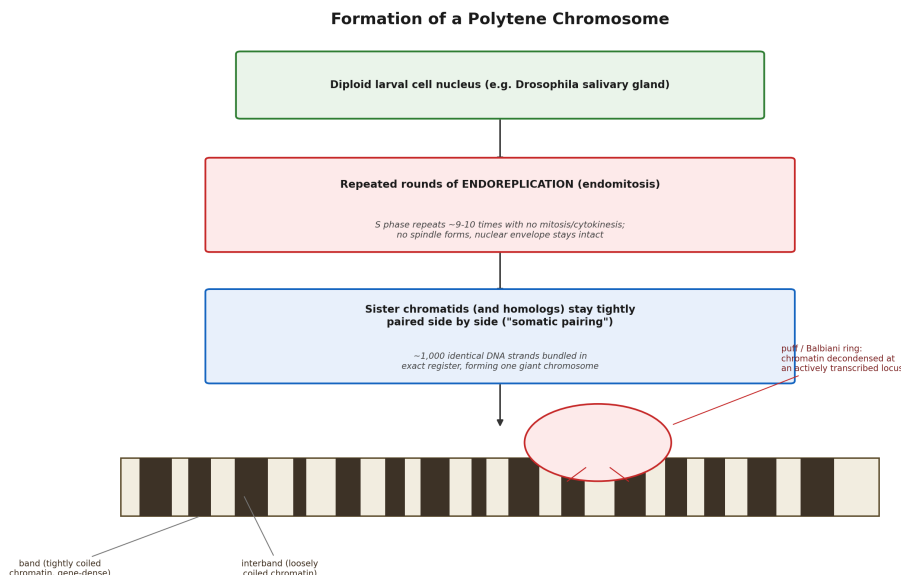


Fig. 2 - Polytene chromosome formation: repeated endoreplication without division, somatic pairing of the resulting strands, and the resulting banded chromosome with a transcriptionally active puff.

2.2 Bands, interbands and puffs

- **Bands:** dark-staining transverse regions where the many aligned DNA strands are tightly coiled and rich in RNA/basic (histone-associated) proteins. Bands are gene-dense and contain about 85% of the chromosome's DNA, despite occupying less physical length than interbands proportionally would suggest.

- **Interbands:** lighter-staining regions between bands where chromatin is more loosely coiled; these contain the remaining ~15% of the DNA and tend to be more accessible to transcription machinery.
- **Puffs:** localised swellings where a band's chromatin decondenses into visible chromatin loops - the cytological signature of a gene (or gene cluster) being actively transcribed at high levels. Puffing patterns change predictably during development (classically demonstrated by ecdysone-hormone-induced puffing sequences), directly linking chromosome structure to gene regulation.
- **Balbiani rings:** especially large puffs, first described in the midge *Chironomus*, representing sites of extremely high transcriptional activity (e.g. genes encoding salivary secretion proteins).

Worth noting

- Because each band/interband pattern is essentially constant and reproducible along a given polytene chromosome, cytologists could - decades before DNA sequencing - construct detailed physical "maps" of *Drosophila* chromosomes and pinpoint the cytological location of specific mutations simply by correlating chromosome rearrangements (deletions, inversions) with altered banding patterns.

3. Sequence Organization of the Eukaryotic Genome

Beyond how chromatin is physically packaged, the underlying DNA *sequence* of a eukaryotic genome is itself organised into distinct classes based on how many times a given sequence is repeated. This was classically revealed by **DNA reassociation (renaturation) kinetics**: genomic DNA is sheared, denatured into single strands, then allowed to reanneal - sequences present in many copies find a complementary partner and reassociate much faster than sequences present as a single copy.

3.1 The Cot curve

Reassociation rate is plotted against **Cot** (pronounced "C-zero-t"): the product of initial single-stranded DNA concentration (C_0) and time (t). Plotting the fraction of DNA reassociated against $\log(\text{Cot})$ produces a characteristic **Cot curve** that resolves a eukaryotic genome into distinct components, each reassociating over a different Cot range depending on its copy number.

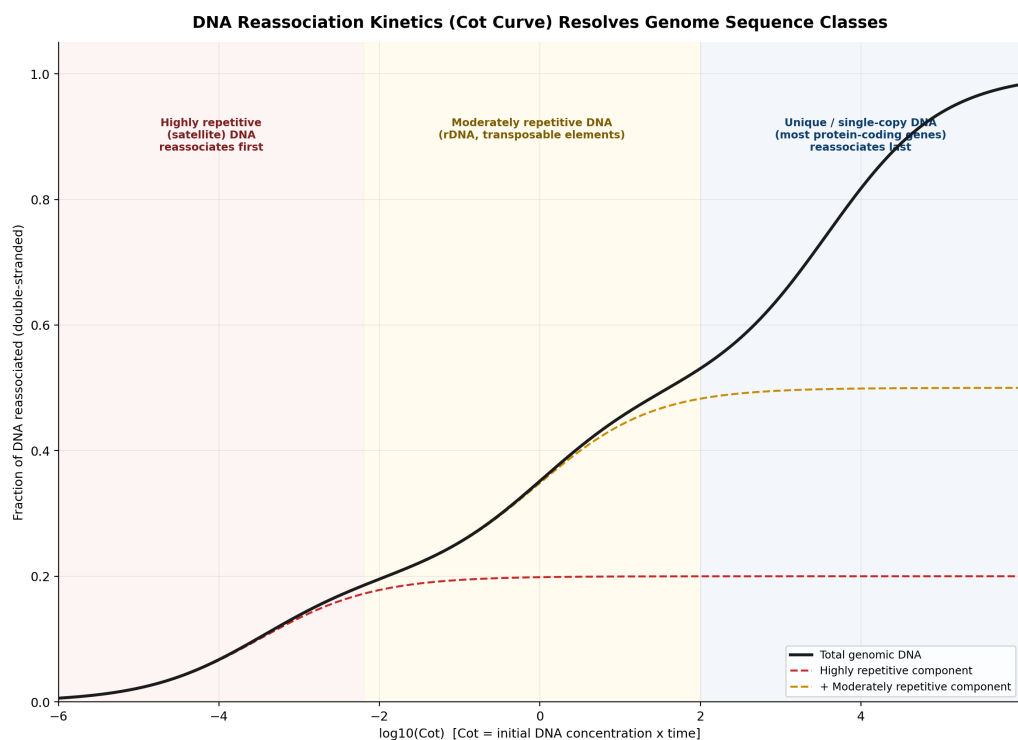


Fig. 3 - A schematic Cot curve: highly repetitive DNA reassociates at low Cot, moderately repetitive DNA at intermediate Cot, and unique/single-copy DNA only at high Cot.

3.2 Three sequence classes

Three Sequence Classes of the Eukaryotic Genome

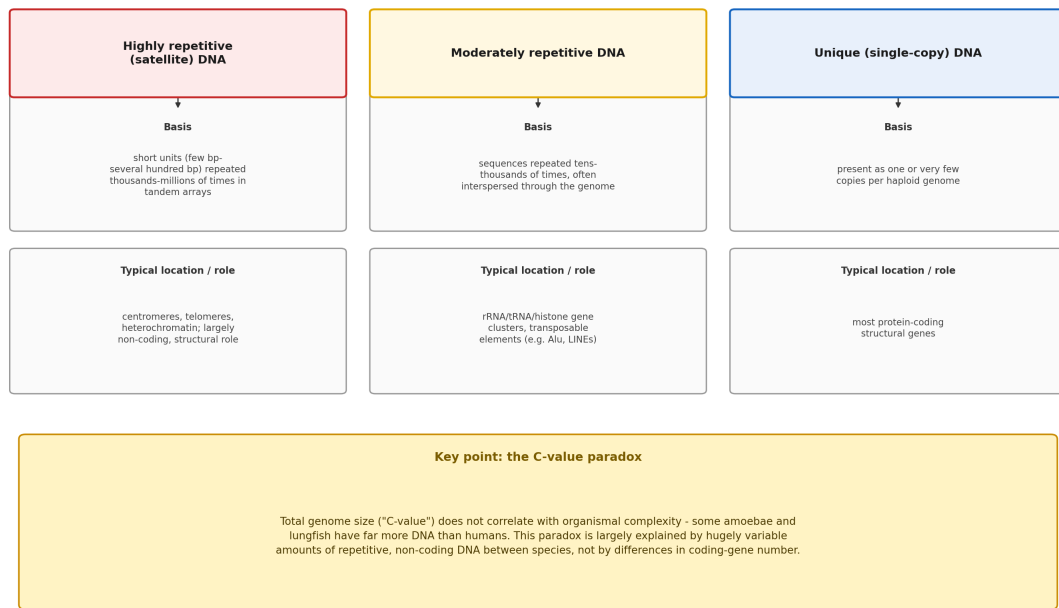


Fig. 4 - The three broad sequence classes of a eukaryotic genome, by basis and typical location/role.

Class	Copy number	Typical examples
Highly repetitive (satellite) DNA	Thousands to millions of copies of a short repeat unit, usually in long tandem arrays	Centromeric and telomeric satellite DNA, constitutive heterochromatin
Moderately repetitive DNA	Tens to thousands of copies, often interspersed throughout the genome rather than tandem	rRNA/tRNA/histone gene clusters; transposable elements (e.g. Alu elements, LINEs)
Unique (single-copy) DNA	One, or very few, copies per haploid genome	Most protein-coding structural genes

3.3 The C-value paradox

The total haploid DNA content of a genome is called its **C-value**. The **C-value paradox** (or C-value enigma) is the observation that genome size correlates poorly with organismal complexity - some amoebae, lungfish and even some plants have vastly more DNA than humans, despite not being obviously more complex. This is now understood to result largely from very large, highly variable amounts of repetitive, non-coding DNA (especially transposable elements) accumulated differently across lineages, rather than from differences in the number of protein-coding genes.

Key points

- Sequence organization (unique vs repetitive) is a separate axis from chromatin packaging (euchromatin vs heterochromatin), though the two are correlated: repetitive DNA is enriched in constitutive heterochromatin, and unique coding DNA is enriched in euchromatin.
- "Repetitive" does not mean "useless" - moderately repetitive rRNA/tRNA gene clusters are essential and heavily used, and transposable elements have contributed to genome evolution and regulatory innovation even though most are not translated into protein.

4. Quick Revision Summary

Concept	Essence
Nucleosome	~146 bp DNA wrapped around a histone octamer; first level of DNA packaging
30 nm fiber	Histone H1-mediated compaction of the nucleosome array
Looped domains / scaffold	30 nm fiber anchored in loops on a non-histone protein scaffold
Centromere / kinetochore	Directs spindle attachment for chromosome segregation
Telomere	Protects and completes replication of chromosome ends
Euchromatin / heterochromatin	Decondensed & active vs condensed & largely silent chromatin
Polytene chromosome	Giant banded chromosome from repeated endoreplication without division
Somatic pairing	Homologs and sister chromatids stay aligned after endoreplication
Bands / interbands	Tightly vs loosely coiled chromatin regions along a polytene chromosome
Puffs / Balbiani rings	Decondensed chromatin loops marking active gene transcription
Cot curve	Reassociation-rate plot resolving genome sequence classes by copy number
Highly / moderately repetitive, unique DNA	Three sequence classes by copy number and typical genomic role
C-value paradox	Genome size does not track organismal complexity - driven by variable repetitive DNA

Notes prepared for study purposes. Consult current primary literature for details required in formal citation or research use.