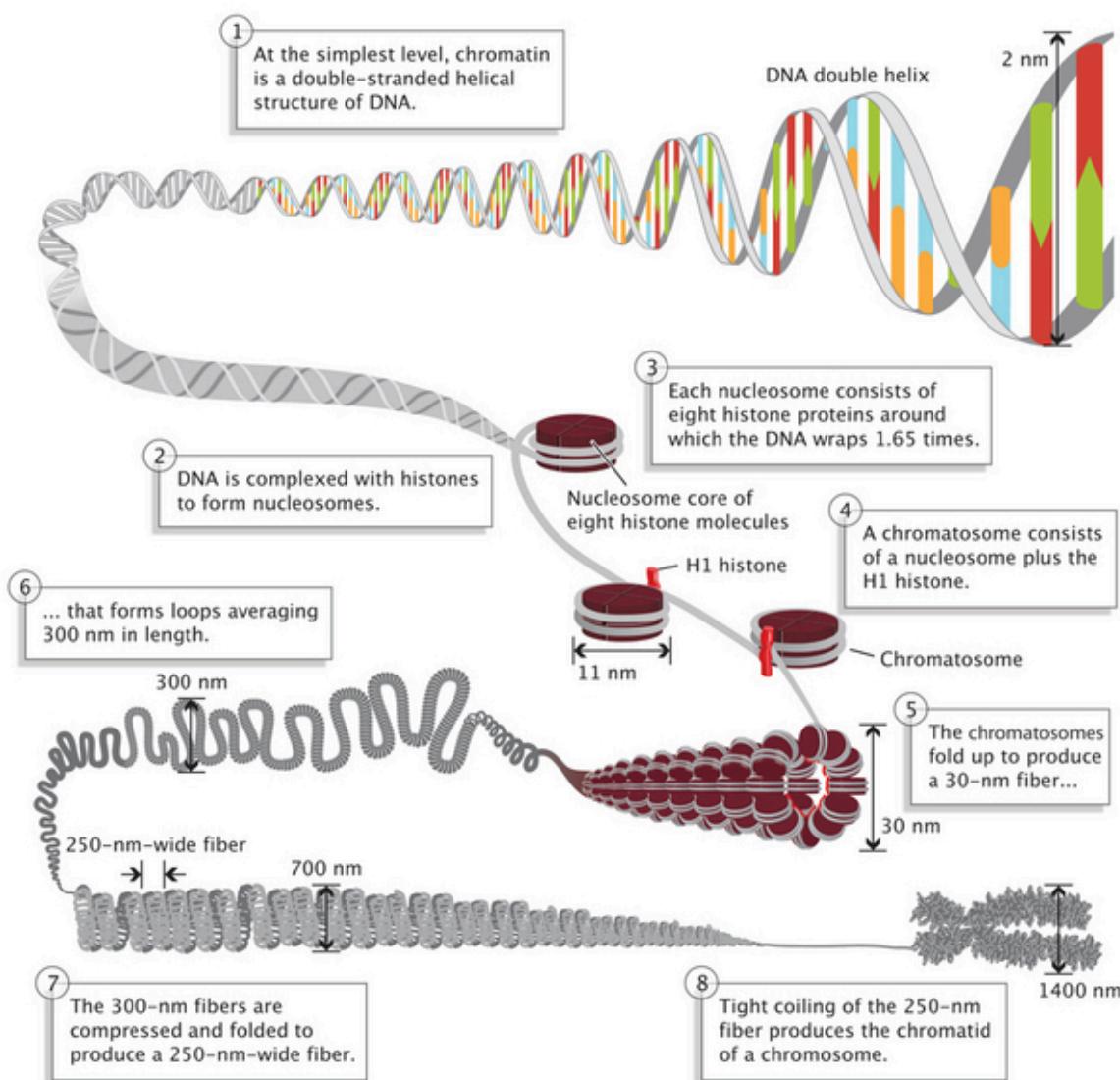


Genome Complexity

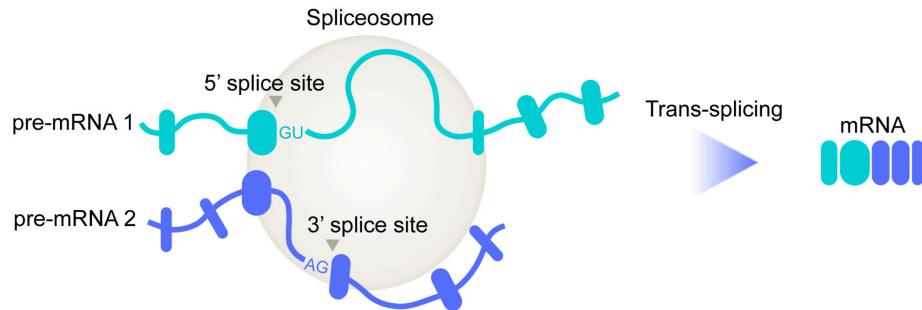
Eukaryotic genomes are linear, unlike bacterial genomes, and follow the Watson-Crick double-helix structure. They are organized into nucleosomes, which are DNA-protein complexes that form chromosomes. While eukaryotic genomes vary in size and gene count, these factors don't necessarily indicate the organism's complexity.

DNA is tightly packed in several steps: it starts as a double helix, wraps around nucleosomes, forms fibers, and eventually compacts into chromosomes. Despite differences in genome size, complexity isn't linked to genome size or gene number.



Having more **protein-coding genes** doesn't always mean an organism is more **complex**. Eukaryotic genomes create complexity in other ways, mainly through how **genes** are **expressed**. One key example is **alternative splicing**, where a single gene can produce **multiple proteins by rearranging its parts** (exons). This means that, even though humans have around 20,000 protein-coding genes, they could produce over 500,000 different proteins.

Other processes, like **RNA editing**, **trans-splicing**, and **tandem chimerism**, also add to genomic complexity. RNA editing changes an mRNA molecule after transcription, sometimes resulting in slight changes to proteins that could be beneficial. Trans-splicing combines different RNA transcripts, and tandem chimerism creates new mRNA molecules by joining nearby genes.



For example, the parasite *Trichomonas vaginalis* has 60,000 protein-coding genes, but most of them don't have introns, so alternative splicing isn't as significant. Many of these genes are duplicates or pseudogenes (non-functional genes), which means their large number doesn't make the organism more complex.

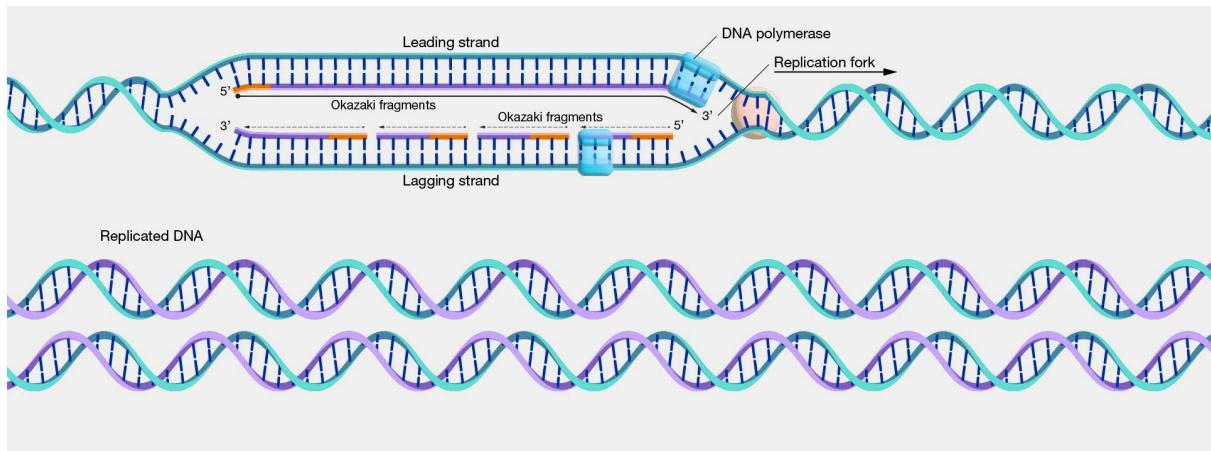
Organism complexity comes from more than just the number of genes or DNA base pairs. It's also about how genes are regulated and expressed, including through noncoding RNA, regulatory elements, and other factors. Thus, **genome size and gene count only contribute a small part to an organism's complexity**.

When scientists refer to the eukaryotic genome, they usually mean the **haploid genome**, which is the full **DNA set in sperm or egg cells**. For humans, this is about 3 billion base pairs. Since **humans have two sets of chromosomes** in most cells, each cell has about **6 billion base pairs**. In humans, the **mitochondrial genome has only about 16,500 base pairs, much smaller than the nuclear genome**.

Table 1: Genome Size and Number of Protein-Coding Genes for a Select Handful of Species

| Species and Common Name | Estimated Total Size of Genome (bp)* | Estimated Number of Protein-Encoding Genes* |
|---|--------------------------------------|---|
| <i>Saccharomyces cerevisiae</i> (unicellular budding yeast) | 12 million | 6,000 |
| <i>Trichomonas vaginalis</i> | 160 million | 60,000 |
| <i>Plasmodium falciparum</i> (unicellular malaria parasite) | 23 million | 5,000 |
| <i>Caenorhabditis elegans</i> (nematode) | 95.5 million | 18,000 |
| <i>Drosophila melanogaster</i> (fruit fly) | 170 million | 14,000 |
| <i>Arabidopsis thaliana</i> (mustard; thale cress) | 125 million | 25,000 |
| <i>Oryza sativa</i> (rice) | 470 million | 51,000 |
| <i>Gallus gallus</i> (chicken) | 1 billion | 20,000-23,000 |
| <i>Canis familiaris</i> (domestic dog) | 2.4 billion | 19,000 |
| <i>Mus musculus</i> (laboratory mouse) | 2.5 billion | 30,000 |
| <i>Homo sapiens</i> (human) | 2.9 billion | 20,000-25,000 |

DNA replication is the process of **copying the DNA** in a cell so that each new cell gets a **complete set of genetic information**. Before a cell divides, it must **duplicate its entire genome**, which contains nearly **three billion DNA base pairs**. DNA polymerases are the molecules that carry out this copying process. It takes several hours to replicate all the DNA in a human cell. After replication, the cell has double the amount of DNA, which is then split between the parent and daughter cells, making them genetically identical.



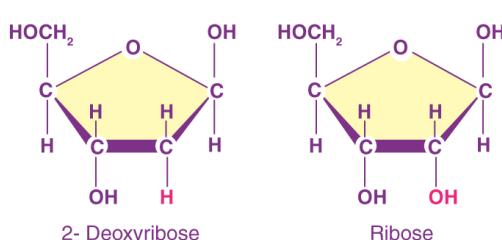
DNA replication is the process by which a **double-stranded DNA molecule is copied to produce two identical molecules**. This ensures that when a cell divides, **both daughter cells have the same genetic information**.

Replication happens in three main steps:

- Separation of Strands:** The DNA double helix uncoils at a specific site called the origin. Proteins like **helicase break the bonds between the DNA strands**, separating them. **A replication fork is a Y-shaped structure** that forms when a DNA double helix splits into two strands.

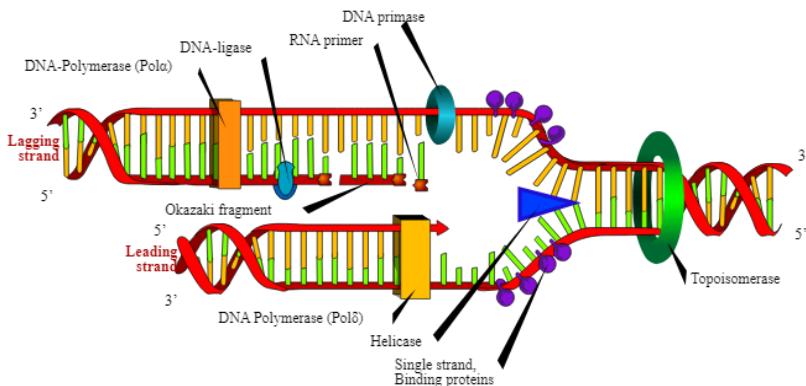


- Priming:** **DNA primase creates short RNA primers** to provide a starting point for DNA polymerase, which builds new DNA strands. Since DNA polymerase can only add nucleotides to an existing strand, the RNA primer serves as a "primer" to kickstart DNA replication on a single-stranded DNA template.

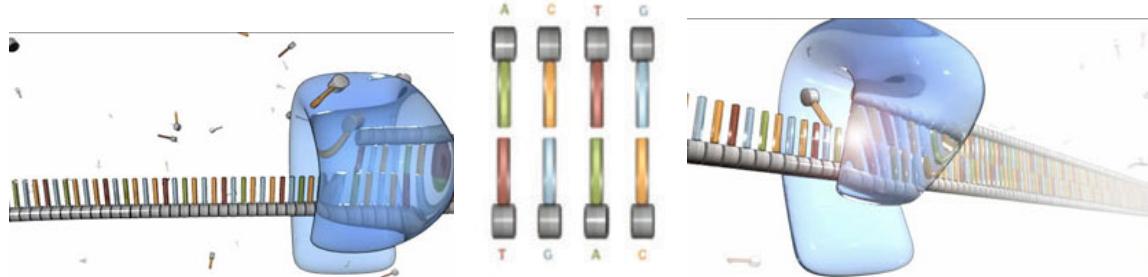


RNA primers are used in DNA replication because DNA polymerase can only add nucleotides to an existing chain with a free 3'-OH group, which RNA primers provide. DNA polymerase can't start DNA synthesis

on its own, so it needs a short RNA primer to begin the process.



3. **DNA Assembly:** The enzyme **DNA polymerase** attaches to the primer and adds new nucleotides to form a complementary strand.



The new strand is created by matching each nucleotide with its complement: **A pairs with T, and C pairs with G.** Replication is fast. In bacteria like *E. coli*, it happens at **1,000 nucleotides per second**, while in humans, it's about **50 nucleotides per second**. The **two strands of DNA are replicated differently**: one (leading strand) is made continuously, while the other (lagging strand) is made in small fragments that are later joined together.

The five DNA polymerases of *E. coli* and some of their relevant properties.

High accuracy in DNA replication is crucial for maintaining genetic stability and preventing harmful mutations. **The error rate during DNA replication is extremely low**, around **1 in a billion to 1 in 100 billion base pairs**. **Achieving this low error rate involves several mechanisms**, especially in the model organism *Escherichia coli* (*E. coli*), where the primary enzyme involved is DNA polymerase III (Pol III).

Here are the key points about **DNA replication fidelity in *E. coli*:**

1. **Pol III is the main enzyme responsible for DNA replication**, but it is supported by other DNA polymerases in the process.
2. **Pol II acts as a back-up proofreader for Pol III**, helping to correct errors made during replication.
3. **Pols IV and V do not significantly contribute to replication accuracy under normal conditions**, but can help with **fidelity if expressed at higher levels**. They primarily function on the lagging strand.

4. Pol I plays a smaller role, mainly in filling in gaps during the formation of Okazaki fragments on the lagging strand. Translesion DNA synthesis (TLS) is a direct mechanism of bypassing unrepaired DNA lesions

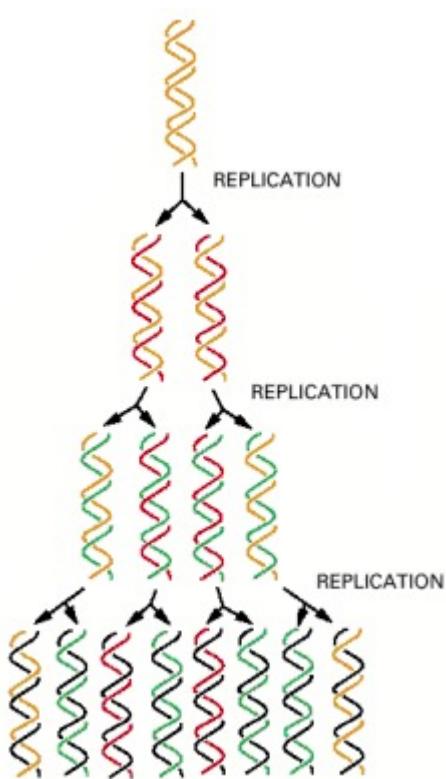
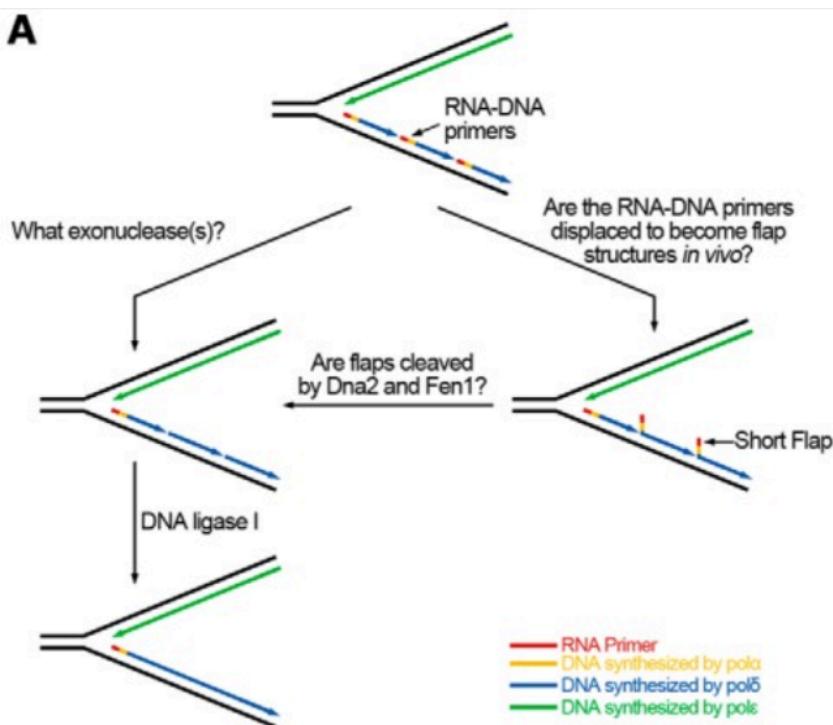
| | Pol I | Pol II | Pol III | Pol IV | Pol V |
|----------------------------------|--|--|--|--------------------|---|
| DNA polymerase family | A | B | C | | |
| Activity | 5'-3' polymerase 3'-5' exonuclease 5'-3' exonuclease | 5'-3' polymerase 3'-5' exonuclease | 5'-3' polymerase 3'-5' exonuclease | 5'-3' polymerase | 5'-3' polymerase |
| | PolA | PolB | Pol III subunits: α, β, δ, ε, τ, γ, ψ, χ | DinB | UmuC/UmuD/UmuD' |
| Number of molecules/cell | - SOS + SOS | 400 400 | 50 - 75 350 - 1000 | 10 - 20 10 - 20 | 150 - 250 1200 - 2500 |
| Biological functions in the cell | DNA replication, Okazaki fragment maturation, DNA repair | DNA replication (backup DNA polymerase), DNA repair, TLS | DNA replication, DNA repair | TLS | TLS |
| | PCNA, Topoisomerase, SSBs, Helicase | Origin | Pol 1 | Parental DNA | 5' Leading strand Okazaki fragments 3' Lagging strand |
| | Nucleotides | Nucleotides | Fork movement | | |

Okazaki fragments are short DNA segments formed **on the lagging strand during replication**. In 1968, Reiji and Tsuneko Okazaki discovered Okazaki fragments while studying bacteriophage DNA replication in *E. coli*. In bacteria and bacteriophage T4, Okazaki fragments are about 1000 to 2000 nucleotides long, while in eukaryotes, they are about 150 to 200 nucleotides long. Each fragment has an RNA primer at the start, and they are later connected by the enzyme DNA ligase to form the lagging strand during DNA replication.

The study investigates how **RNA-DNA primers are removed from Okazaki fragments during DNA replication, focusing on different pathways**. There are two proposed ways to remove these primers: **the exonuclease pathway and the flap pathway**. In the exonuclease pathway, the RNA-DNA primers are directly digested by enzymes like RNase H2 and Exo1. In the flap pathway, the RNA-DNA primers are first displaced, creating "flap" structures. These flaps are then cut by endonucleases like Fen1 and Dna2.

<https://www.youtube.com/watch?v=TNKWgcFPHqw>

Exonucleases are enzymes that break down DNA by removing nucleotides from the ends of DNA strands. They are crucial for maintaining genome stability and play a role in DNA replication, repair, and cell metabolism. Exonucleolytic proofreading by DNA polymerase during DNA replication.



The semiconservative nature of DNA replication. In a round of replication, each of the two strands of DNA is used as a template for the formation of a complementary DNA strand. The original strands therefore remain intact through many cell generations.

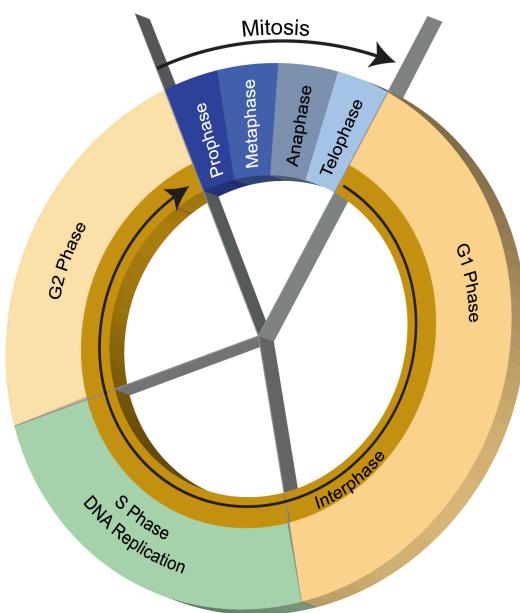
A human chromosome has about 150 million nucleotide pairs. Replicating this DNA with a single replication fork moving at 50 nucleotides per second would take about 800 hours. To speed up the process, many replication forks work at the same time on each chromosome. These forks are grouped in clusters called replication units, with 20-80 origins. Replication starts at different times during the cell cycle until all DNA is copied. Within a replication unit, origins are spaced 30,000–300,000 nucleotide pairs apart. Replication forks move in opposite directions, forming a bubble and stopping when they meet other forks or the end of the chromosome. This allows many forks to replicate the DNA quickly and efficiently.

| Topic | Prokaryotic DNA Replication | Eukaryotic DNA Replication |
|---|--|---|
| Definition | Prokaryotic DNA replication is the process by which a prokaryotic organism duplicates its entire genome in order to pass the second copy to a daughter cell. | Eukaryotic DNA replication is the process by which the eukaryotic genome duplicates prior to cell division. |
| Occurrence | Prokaryotic DNA replication is a continuous process. | Eukaryotic DNA replication occurs during the S phase of the cell cycle. |
| Location | Prokaryotic DNA replication takes place in the cytoplasm. | Eukaryotic DNA replication takes place in the nucleus. |
| Type of DNA | Prokaryotic DNA is circular and double-stranded. | Eukaryotic DNA is linear and double-stranded with ends. |
| Amount of DNA | There is a small amount of Prokaryotic DNA. | The amount of eukaryotic DNA is 50 times more than the amount of prokaryotic DNA. |
| Packaging | Prokaryotic DNA forms loop-like structures by wrapping around histone-like protein molecules. | Eukaryotic DNA forms nucleosomes and shows higher order packaging. |
| Origin of Replication | Prokaryotic DNA consists of a single origin of replication. | Eukaryotic DNA consists of multiple origins of replication (over 1000). |
| DNA Polymerases | Prokaryotic DNA replication is carried out by DNA polymerase I and III. | Eukaryotic DNA replication is carried by DNA polymerase α , δ , and ϵ . |
| Size of the Okazaki Fragment | The Okazaki fragments are comparatively large, 1000-2000 nucleotides in length. | The Okazaki fragments are small, around 100-200 nucleotides in length. |
| DNA Gyrase | DNA gyrase is involved in the prokaryotic DNA replication. | DNA gyrase is not required for the eukaryotic DNA replication. |
| Rate of DNA replication | Prokaryotic DNA replication is a rapid process and around 2000 nucleotides are added per second. | Eukaryotic DNA replication is a slow process and around 100 nucleotides are added per second. |
| End Synthesis | Prokaryotic DNA does not contain ends. | Telomerase is involved in the end synthesis in Eukaryotic DNA during the replication. |
| Final Product of the Replication | The final product of the prokaryotic DNA replication is two circular chromosomes. | The final product of the eukaryotic DNA replication is two sister chromatids. |

CELL CYCLE

The cell cycle is the process through which a **cell replicates and forms two new cells**. It has **four stages: G1/(Gap/Growth), S, G2, and M**.

In **G1**, the cell prepares to divide. The cell prepares for division. At a point called the **restriction point**, the cell commits to division and enters **S phase**.



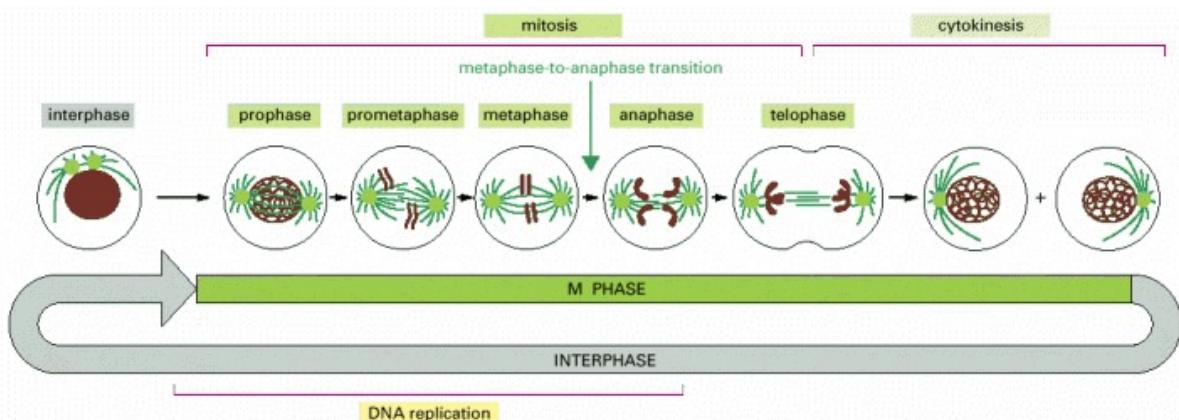
In the **S phase**, the cell copies its DNA (S stands for DNA synthesis). DNA is copied, and each chromosome now has two sister chromatids.

After that, in **G2**, the cell prepares for mitosis and cytokinesis by assembling the necessary materials.

The longer part of the cycle is **interphase**. Mitosis consists of four stages, with a significant change in the cell's biochemical state happening when it moves from metaphase to anaphase. The cell can pause in metaphase, but once it passes this point, it continues through mitosis, cytokinesis, and into interphase. **DNA replication takes place during interphase.**

membrane then reforms, and the chromosomes begin to decondense into their interphase conformations. **Telophase is followed by cytokinesis, Once M is complete, the cell divides into two, and the cycle starts over.** In the **M phase**, the cell divides the DNA into two parts during mitosis. **Mitosis (nuclear division) and cytokinesis (cell division)**, together known as M phase, make up only a small portion of the cell cycle.

https://www.youtube.com/watch?v=5bq1To_RKEo

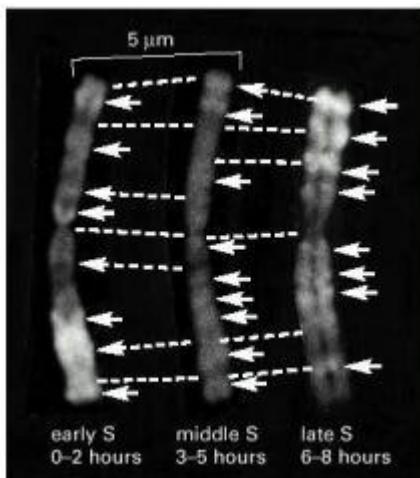


In mammalian cells, DNA replication in the region between two replication origins typically takes about an hour, but the S phase itself lasts around 8 hours. This suggests that replication origins are activated at different times, and each replication unit is only replicated during a small portion of the total S-phase.

The timing of DNA replication is influenced by the structure of the chromatin in which the replication origins are located. Heterochromatin, a tightly packed form of chromatin, replicates late in S phase, while transcriptionally active chromatin, which is less condensed, replicates earlier.

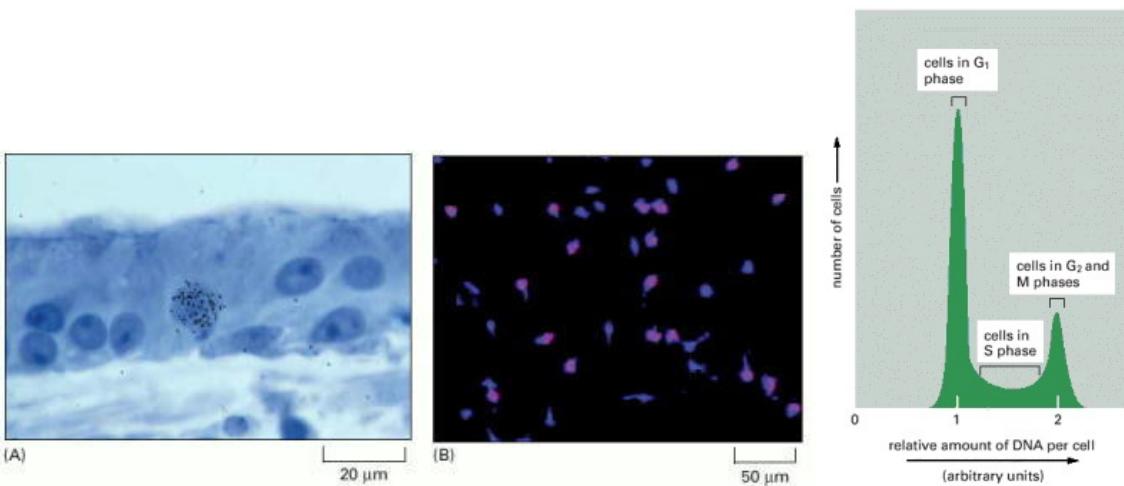
Cell-cycle progression can be studied in several ways. One method is by observing cells under a microscope. Some cells will be in mitosis or cytokinesis, while cells in S-phase (DNA synthesis) are harder to detect. However, S-phase cells can be identified by

incorporating labeled molecules like 3H-thymidine or BrdU into their newly synthesized DNA. These labeled cells can then be visualized using autoradiography or anti-BrdU antibodies.



In a rapidly proliferating cell population, about 30-40% of cells will be in S phase at any given time. By calculating the proportion of labeled cells, the duration of S phase can be estimated. Similarly, the proportion of cells in mitosis (mitotic index) helps estimate the duration of M phase. By tracking labeled cells over time, one can also determine how long it takes for a cell to progress through different phases of the cycle.

Another method to assess cell-cycle progression is by measuring DNA content, which doubles during S phase. Flow cytometry, using DNA-binding fluorescent dyes, allows rapid measurement of DNA content across many cells. This method can also help determine the duration of G1, S, and G2 + M phases by tracking a synchronized population of cells.

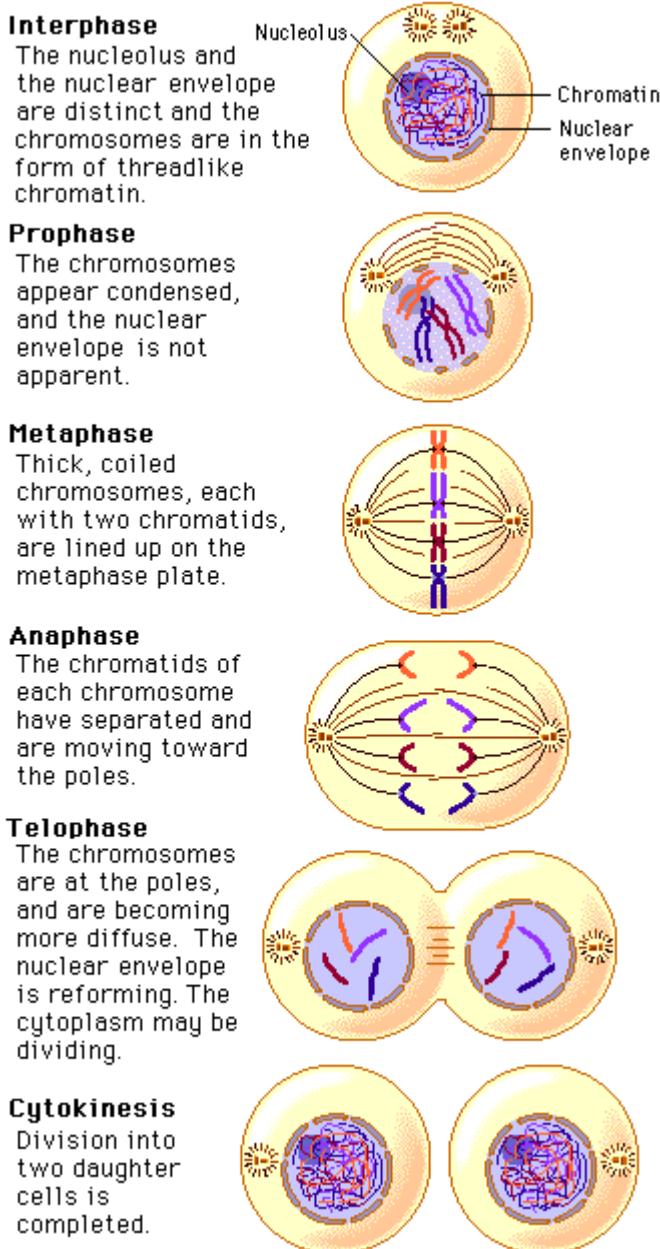


Flow cytometry:

The cells were stained with a dye that fluoresces when bound to DNA, so the brightness indicates the amount of DNA in each cell. The cells are categorized into three groups: G1 phase (unreplicated DNA), G2 or M phase (fully replicated DNA, double the G1 amount), and S phase (DNA replication in progress, intermediate DNA content). **The graph shows that most cells are in G1 phase**, indicating G1 lasts longer than G2 + M phases in this population.



<https://www.youtube.com/watch?v=xPk-kVTUH5c>



Mitosis does not happen in prokaryotes; they divide through a simpler process called binary fission. This difference is mainly due to prokaryotes lacking a nucleus, so they don't need mitosis.

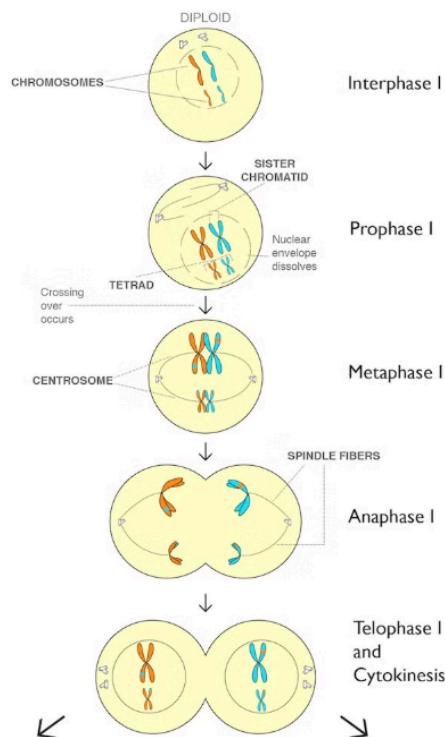
MEIOSIS

<https://www.youtube.com/watch?v=kQu6Yfrr6j0>

Meiosis is the process by which sexually reproducing organisms create gametes (sperm and eggs). It is crucial for forming embryos, ensuring genetic diversity, and preparing for reproduction. The main function of meiosis is to reduce the number of chromosomes from diploid (2 sets) to haploid (1 set). This reduction is vital for successful fertilization and normal development. Meiosis can also cause genetic errors or birth defects, such as Down syndrome, if problems occur. Meiosis occurs in the

primordial germ cells, cells specified for sexual reproduction and separate from the body's normal somatic cells.

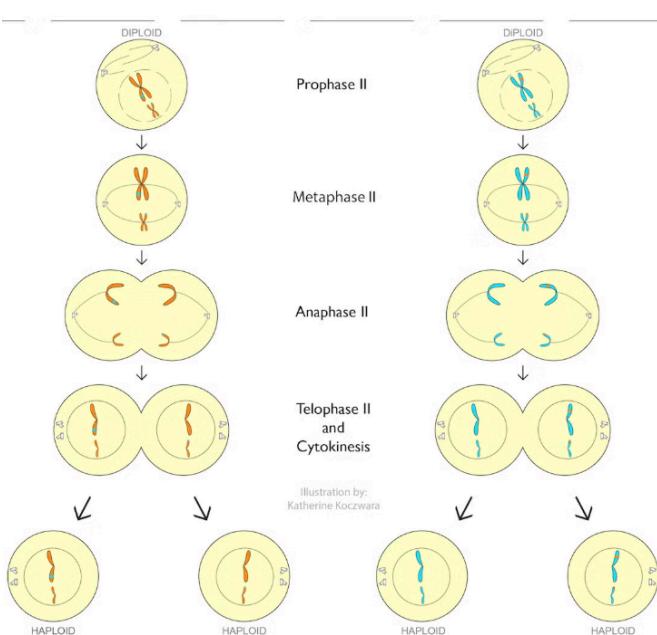
- Meiosis was first described in the 1870s and is essential for creating gametes in both males (spermatogenesis) and females (oogenesis). It starts in primordial germ cells, which are distinct from regular body cells. **These cells undergo DNA replication during interphase before meiosis begins.**



- Meiosis I: The number of chromosomes is halved, but the number of cells doubles.**

- Prophase I (2N):** Chromosomes coil, the nuclear membrane breaks down, and centrosomes move apart. Crossing over occurs, where chromosomes exchange pieces.
- Metaphase I (2N):** Pairs of chromosomes (bivalents) align in the center, and spindle fibers attach.
- Anaphase I:** Homologous chromosomes (one from each parent) are pulled apart, resulting in $N + N$.
- Telophase I:** The nuclear envelope reforms, and the cell splits into two $N + N$ daughter cells, each with half the original chromosome number.

- Meiosis II:** Similar to mitosis, where chromosomes are not reduced but divided into separate cells (haploid cells)



- Prophase II ($N + N$):** Chromosomes coil, and the nuclear membrane breaks down.

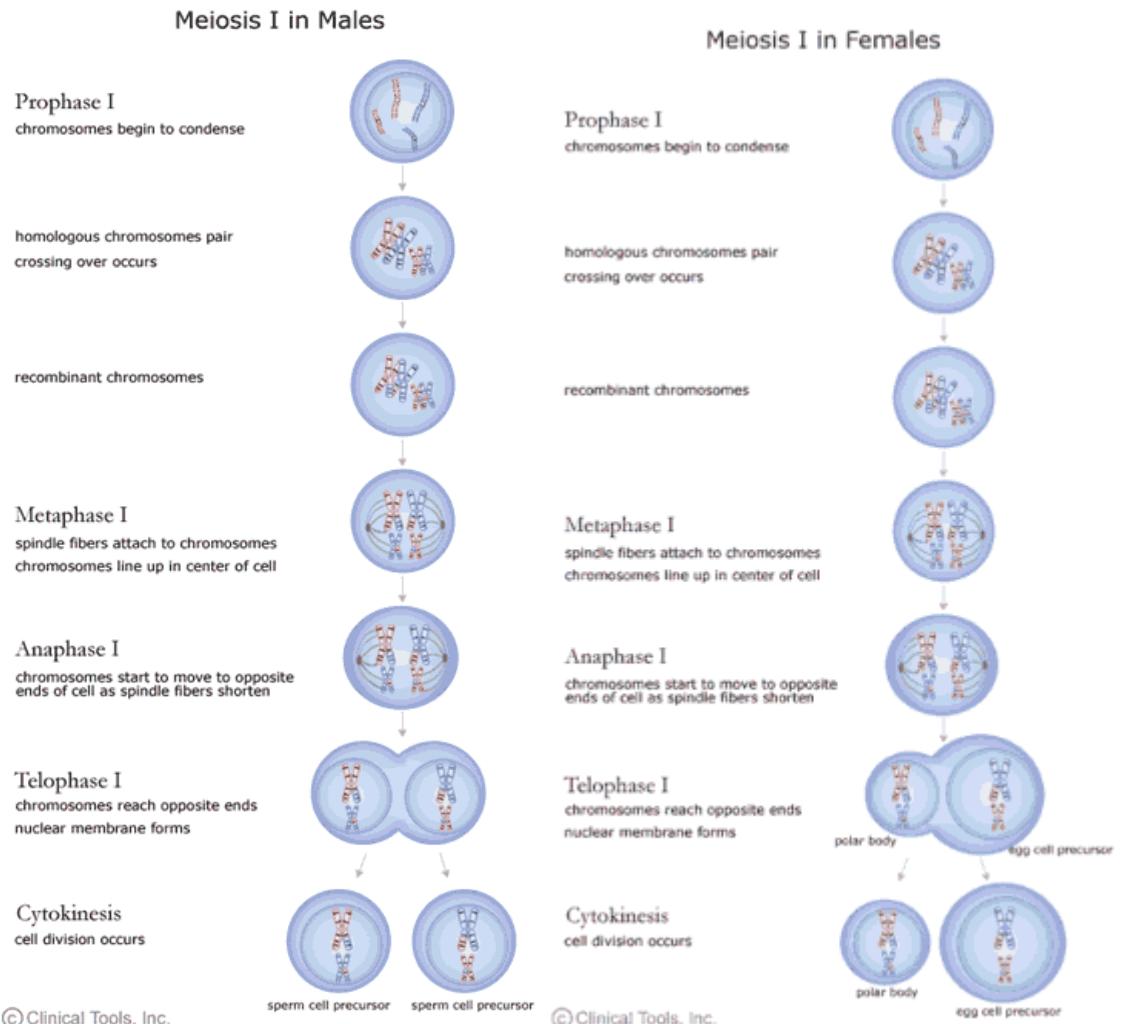
- Metaphase II ($N + N$):** Spindle fibers form, and sister chromatids align along the equator, each with $N + N$ chromosomes.

- Anaphase II:** Sister chromatids separate, resulting in $N + N + N + N$.

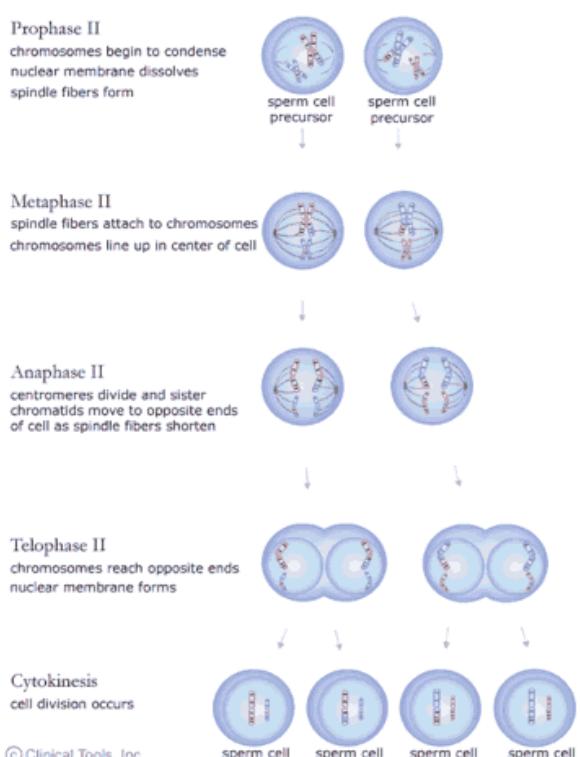
- Telophase II & Cytokinesis:** The chromatids reach the poles and uncoil into chromatin, and the nuclear membrane reforms. The result is four N haploid daughter cells.

- Meiosis occurs in different times and different locations depending on the sex
- Spermatogenesis (in males) produces four sperm cells from one germ cell and occurs continuously after puberty.
- Oogenesis (in females) produces one egg and smaller polar bodies, due to unequal division of cytoplasm. Occurs mostly during fetal development and later in life.

Meiosis ensures genetic integrity and diversity by reducing chromosome numbers. Errors can result in conditions like Down syndrome. This process is conserved in all sexually reproducing organisms, highlighting its importance in reproduction and evolution.



Meiosis II in Males



Meiosis II in Females

