



# Molecular Biology (4)

## DNA replication

Prof. Mamoun Ahram  
School of Medicine  
Second year, Second semester, 2024-2025

#Not edited

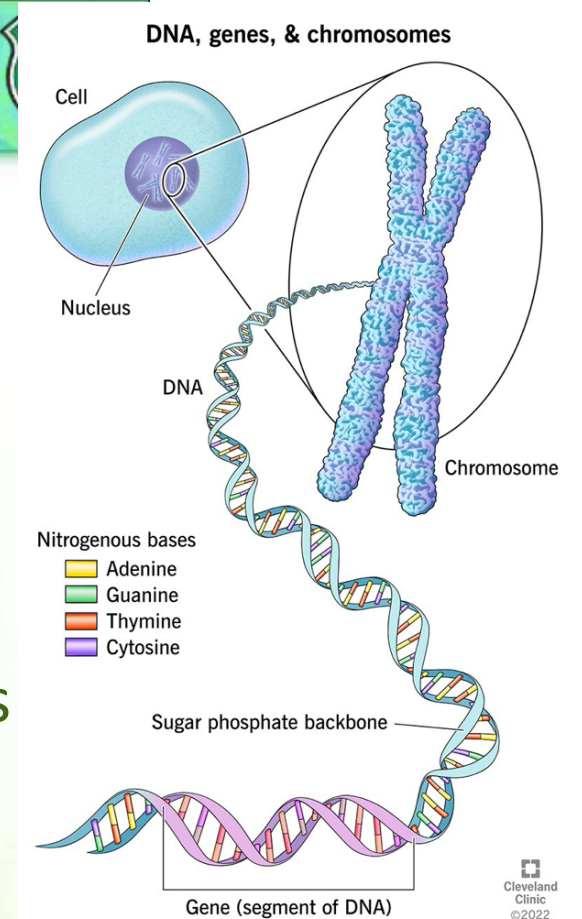
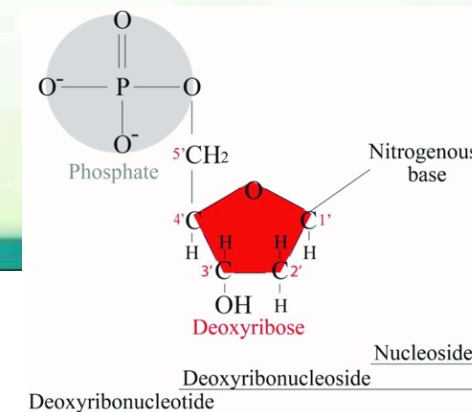
اصْنَعْنَا لِلْمَيْدَانِ يَا رَبِّ.. لَا تَجْعَلْنَا مَمَّنْ يُطِيلُ وَقُوفًا عَلَى فِرَاحٍ، ارزقنا عُكُوفًا صَادِقًا عَلَى ثَغْرِ تَحِيَّهِ، وَإِنْ لَمْ يَرَهُ أَحَدٌ، رَبَّنَا فِي الظِّلِّ،  
عُدْنَا لِلنَّصْرِ، أَلْهَمْنَا الصَّبْرَ، وَاَنْصُرْنَا فِي مَعَارِكِ النَّفْسِ الْخَفِيَّةِ، وَشِدَائِدِ الدَّرْبِ الْقَوِيَّةِ، اجْعَلْنَا جُنُودًا لَكَ، قَلُوبًا فِي السَّمَاءِ. ♥

# Some basic information

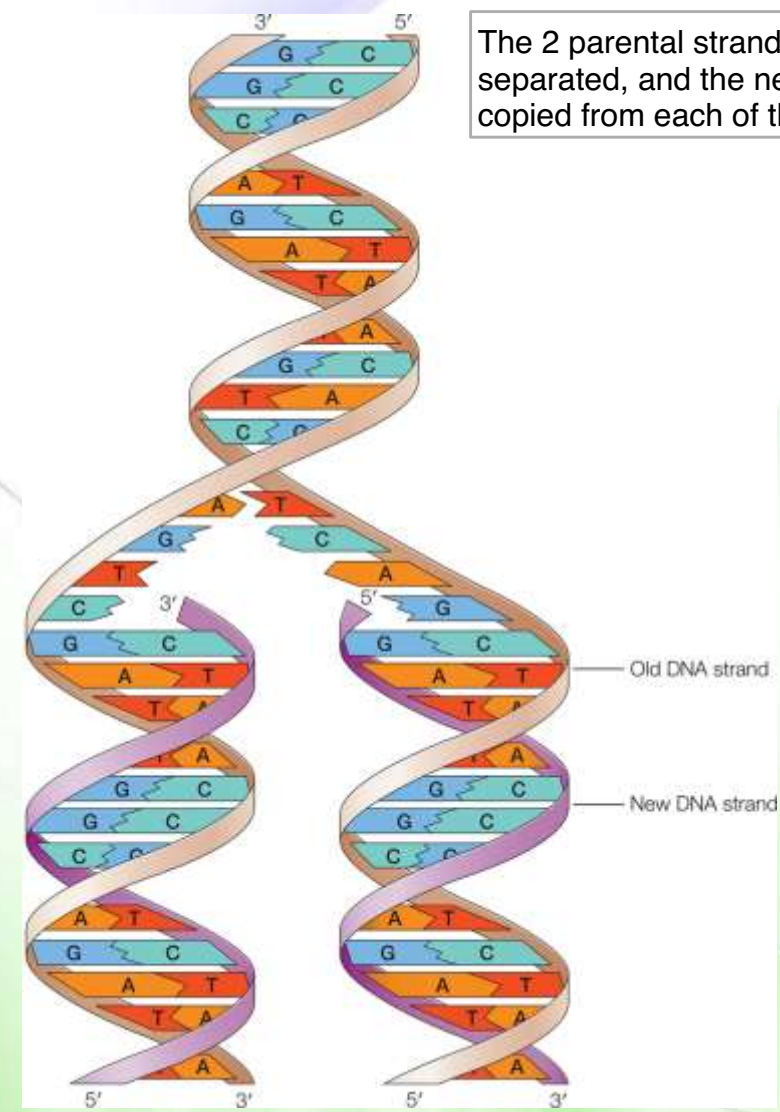
- The entire DNA content of the cell (or an organism) is known as a “genome”.
- DNA is organized into **chromosomes**.
  - Bacterial genome: usually one, circular chromosome.
  - Eukaryotic genome: multiple, linear chromosomes complexed with proteins known as **histones**, and the complex is known as **chromatin**.
- DNA must be accurately **copied (replicated)**.
- DNA **synthesis** is carried out by **DNA polymerases**.
  - In bacteria (E. coli: DNA polymerases I, II, and III)
  - In Humans (DNA polymerases  $\alpha$ ,  $\delta$ , and  $\epsilon$ )
- The substrates are **deoxyribonucleotides**.

Or deoxyribonucleoside triphosphate

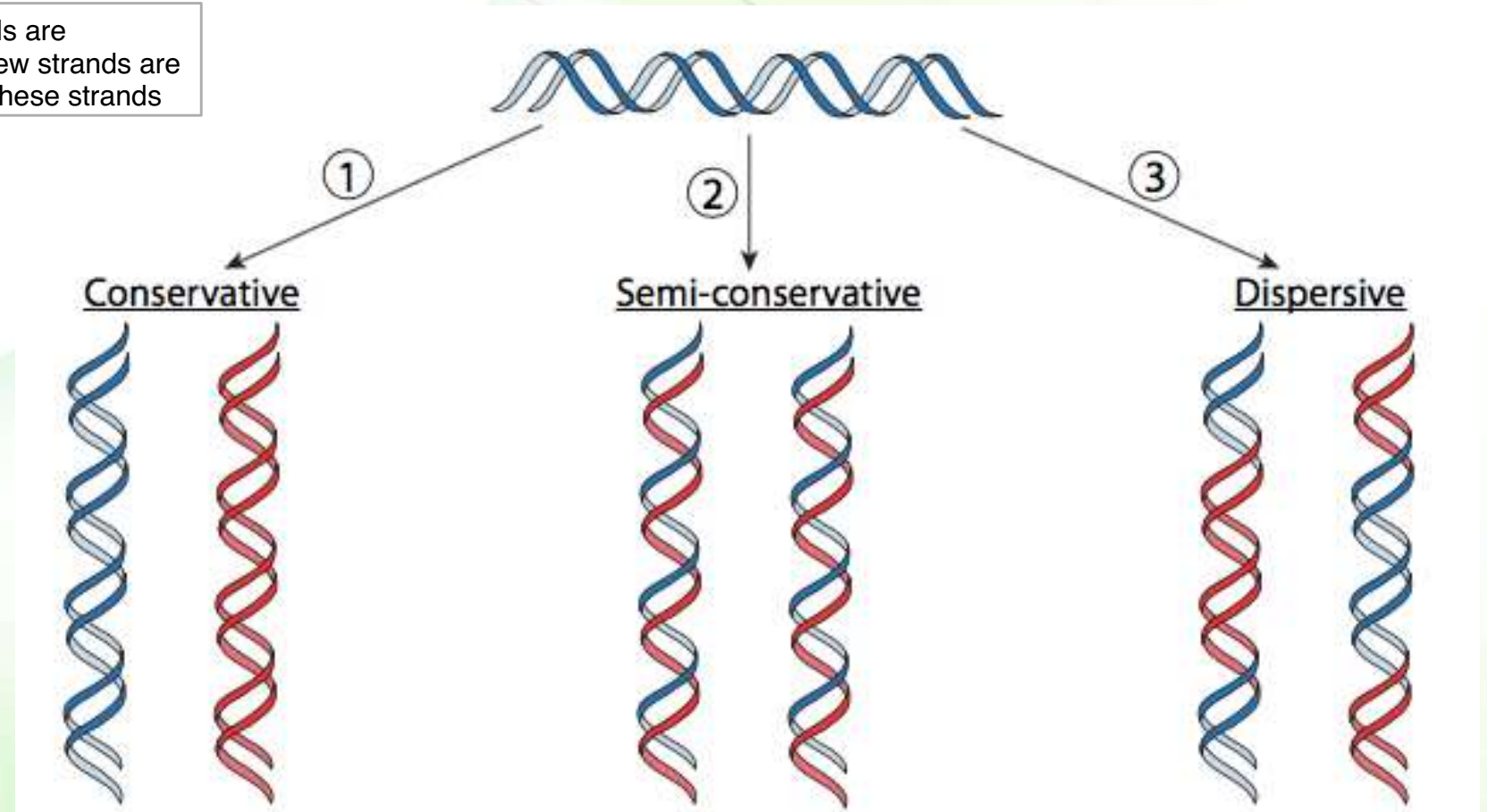
Alpha, delta and epsilon



# The hypotheses and fact



The 2 parental strands are separated, and the new strands are copied from each of these strands



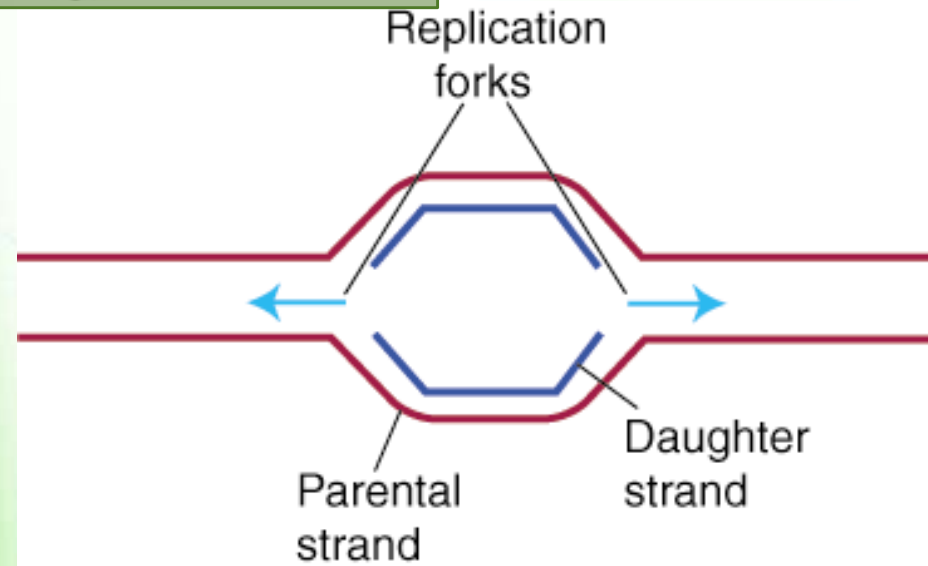
3 conservative hypotheses:-  
1) conservative:- the old DNA is totally (completely) conserved  
2) occurs randomly (as fragments) (the new DNA will consist of old dsDNA and new dsDNA  
3) semi-conservative:- the daughter cells will have DNA consisting of old strand+new strand

# Bidirectional

- Replication is bidirectional.
- This replicative region is called a replication fork.

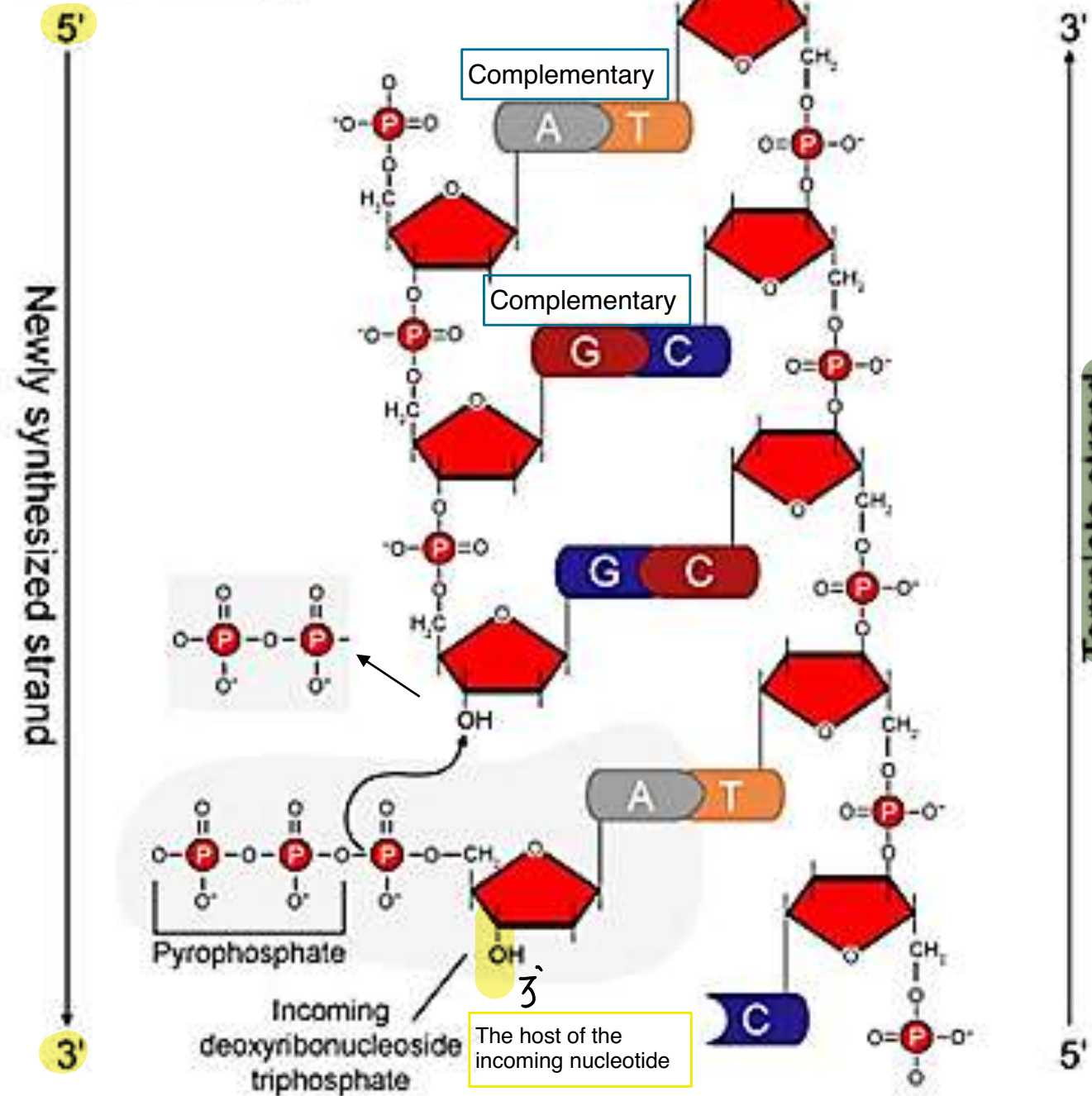


DNA replication goes in 2 opposite directions, left and right



## DNA Synthesis

Elongation from 5' → 3'



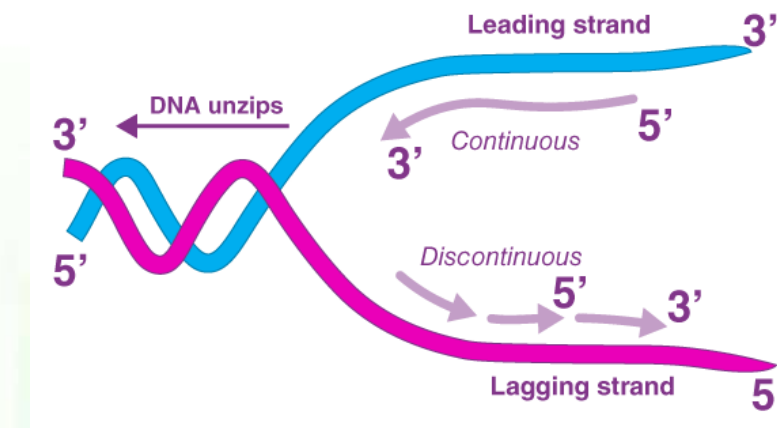
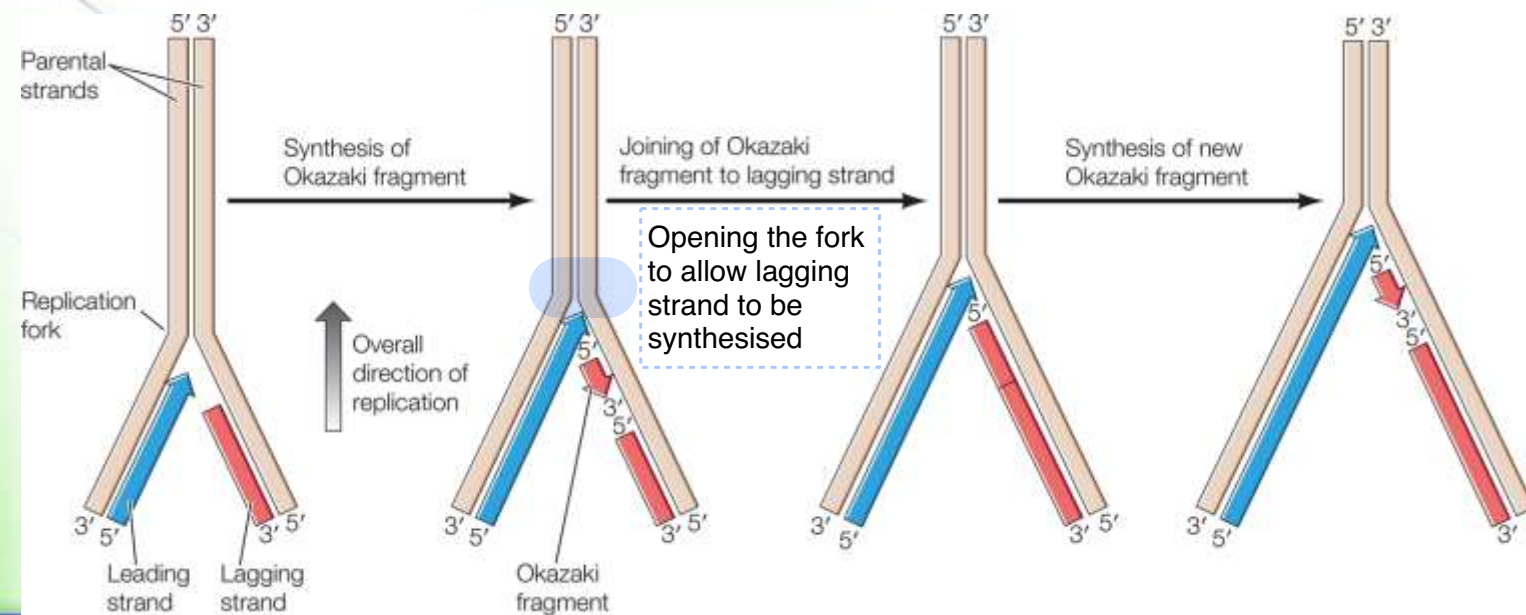
Deoxyribonucleoside triphosphate is the **substrate**, it's added to the 3' end of the newly synthesised strand, after that 2 phosphates are released, so you end up with a nucleotide that has just one phosphate, and this phosphate connects the nucleotides to each other (phosphodiester bond). Because it's building up, we need energy, the energy comes from the substrate itself.

The original DNA strand

# Continuity of DNA synthesis



- The parental DNA is called **a template**.
- The new DNA is synthesized **ONLY** from the 5'-end to the 3'-end.
- One strand of DNA is **continuously** synthesized and called the leading strand.
- The other strand is synthesized **discontinuously** as shorter pieces known as Okazaki fragments and is called the lagging strand.



5

It's called lagging strand ;  
Because it's synthesis is lagging behind(not completed until the leading strand allows it to be synthesised)  
=> the leading strand synthesised continuously,opening up the fork further,allowing the synthesis of the lagging strand, so it waits the leading strand to open up the fork, and it's synthesised as short fragments (Okazaki fragments) which then been connected together by **ligase enzyme**



# Components of DNA replication

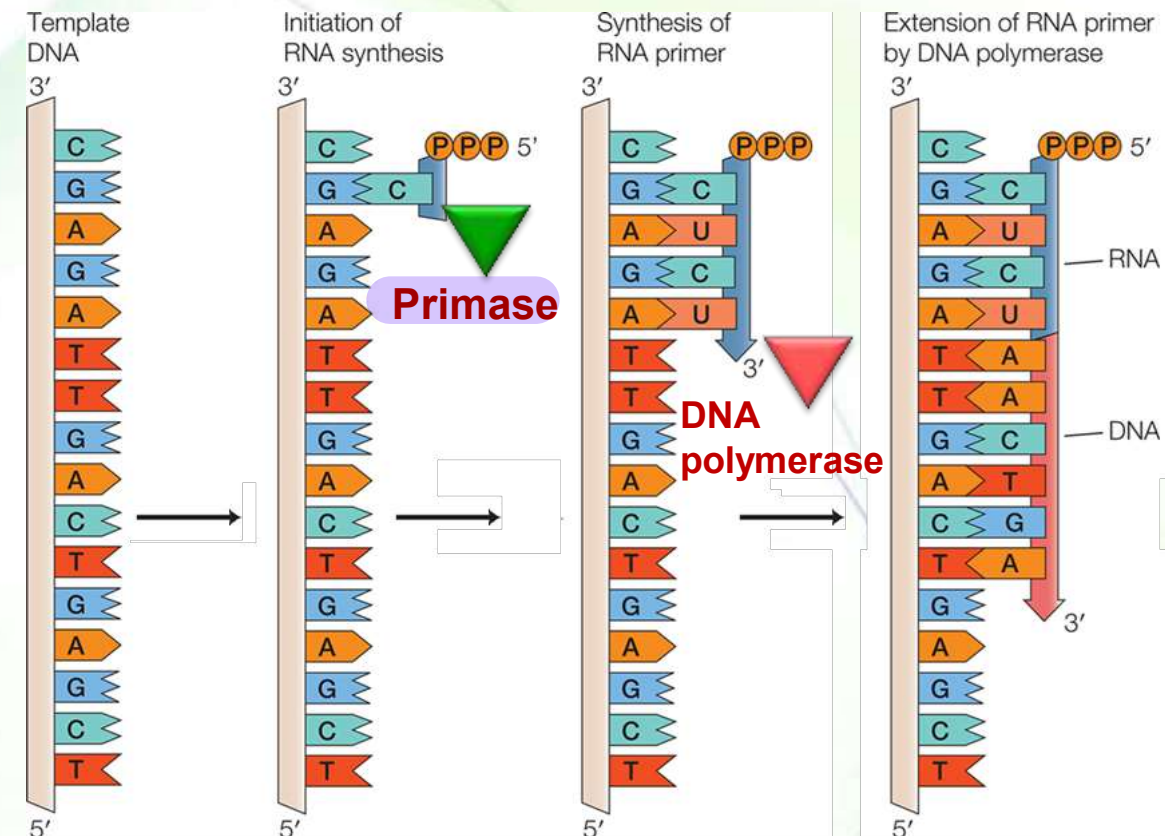
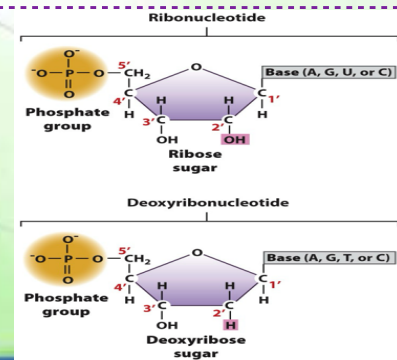
# Primase and RNA primer



It can't set on DNA and start synthesis by itself

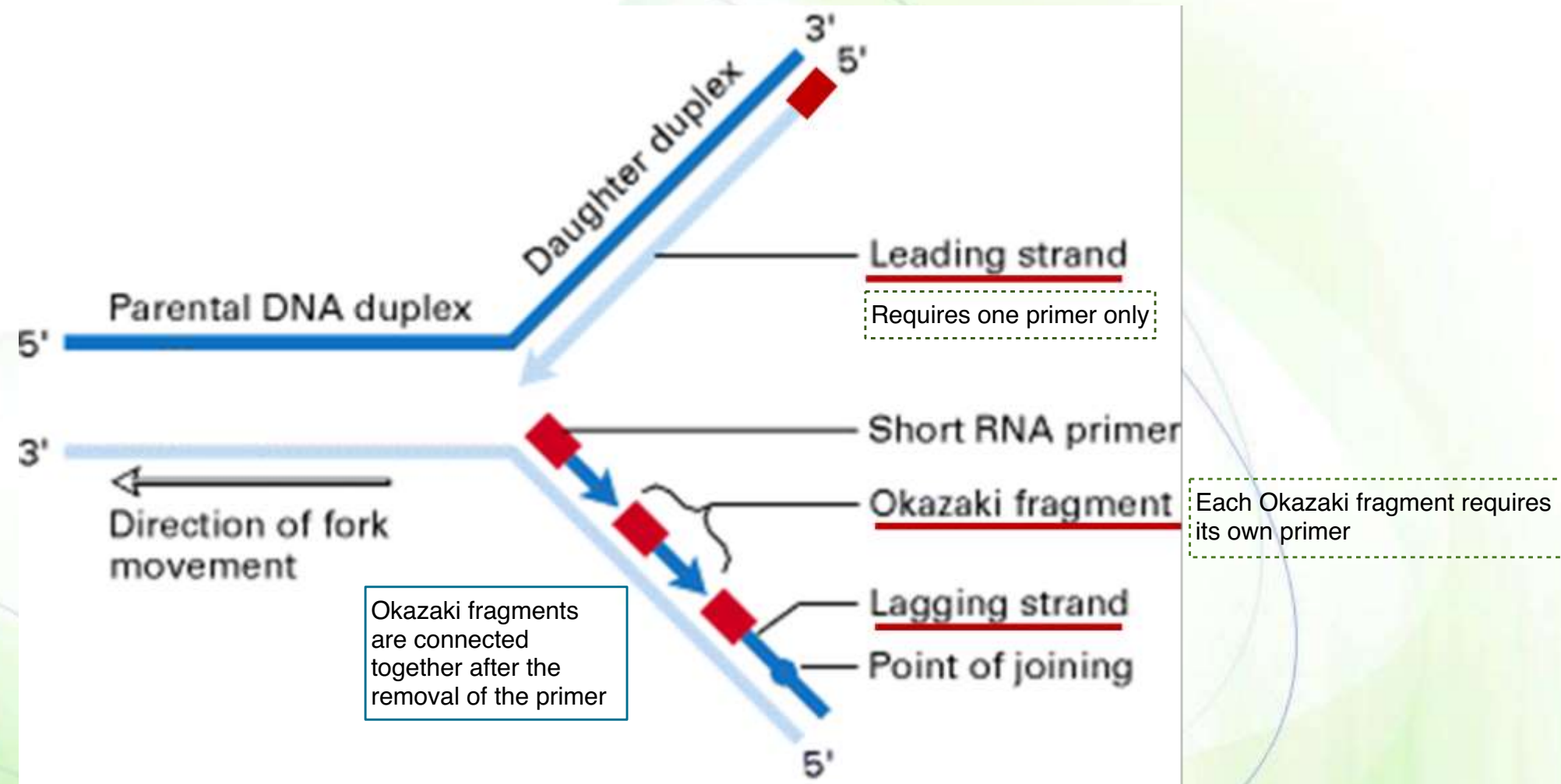
- DNA polymerases cannot initiate replication *de novo* (from scratch).
- They require a RNA primer (3-10 nucleotides long) that is complementary to the DNA template to be added first.
- It is synthesized by a primase.

Except that you have **U** instead of T & **ribonucleoside triphosphate** as a substrate instead of deoxyribonucleoside triphosphate



- 1) RNA primer(fragment) synthesized firstly
- 2) DNA polymerase comes then to start synthesis by adding new nucleotides using the template

# The need of primers

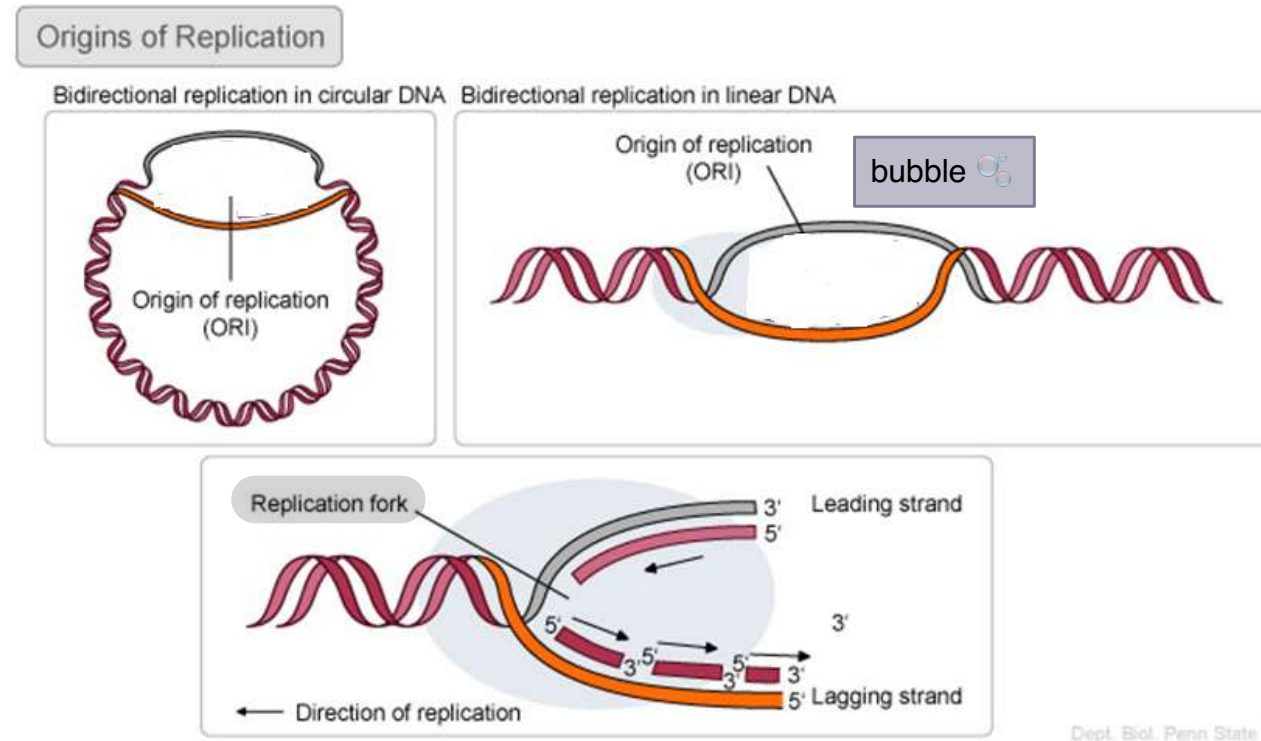




# Exercise



- I have shown you how DNA synthesis proceeds in the replication fork. Draw how DNA replication proceeds in the whole bubble.



Extra:-

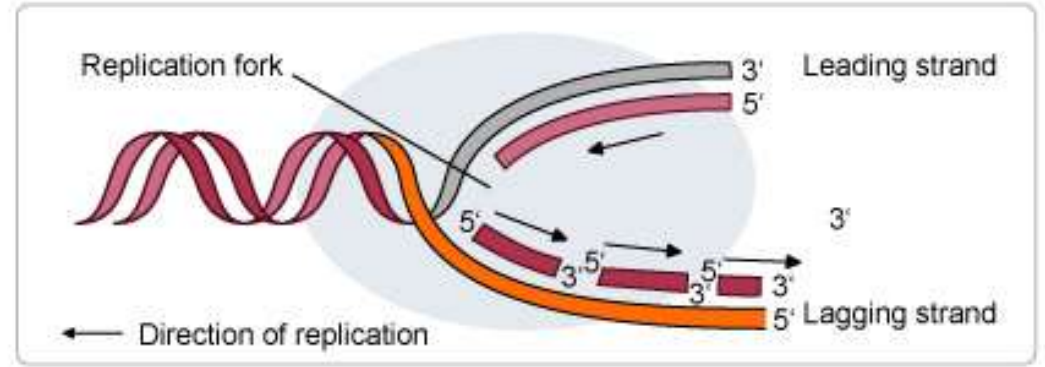
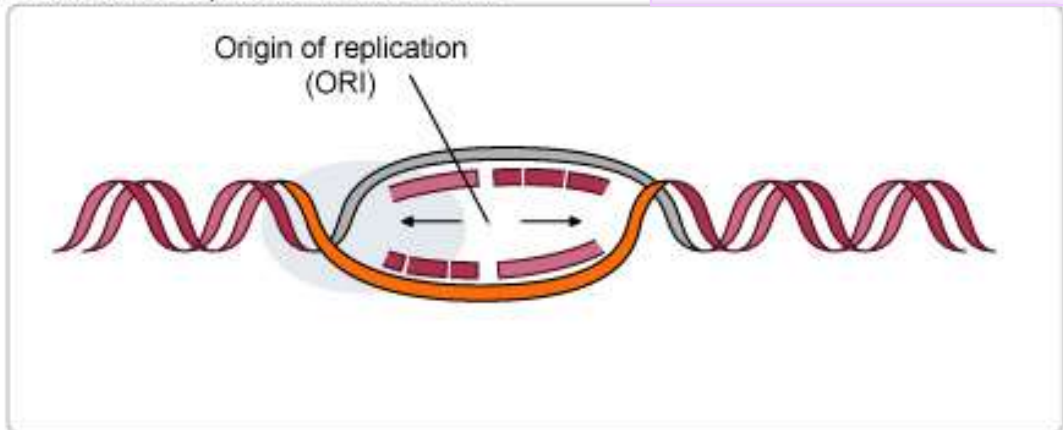
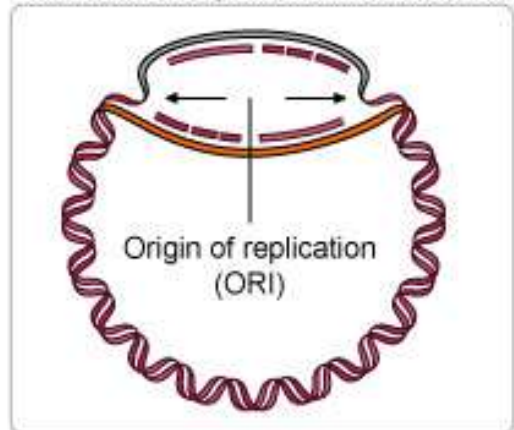
\*\* DNA unwinding at the ORI forms a structure called the replication **bubble**.

\*\* The bubble expands as replication proceeds, with replication occurring at both ends of the bubble (replication forks).



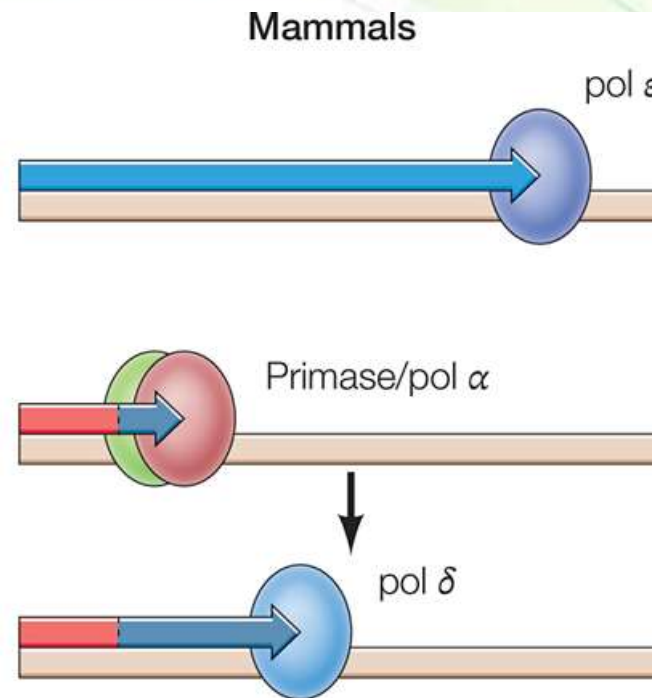
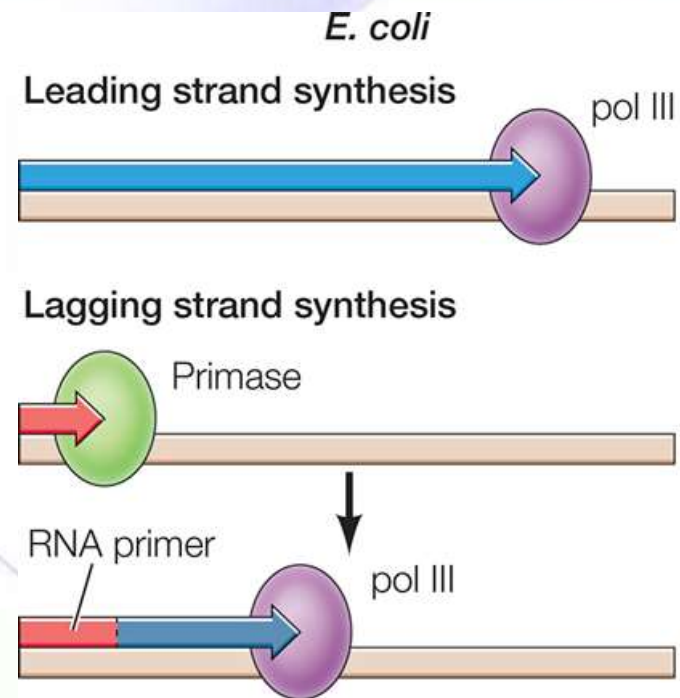
## Origins of Replication

Bidirectional replication in circular DNA    Bidirectional replication in linear DNA



Extra:-  
Replication starts at the origin of replication (ORI) and proceeds in two directions, forming a replication bubble.  
Each replication bubble has two replication forks moving in opposite directions.

# The replicative process



- 1) primase associates polymerase alpha, they set on DNA, Primase synthesize RNA primer then DNA polymerase **alpha** synthesises the first portion of DNA
- 2) in **lagging strand**, polymerase **delta** continues DNA synthesis
- 3) in **leading strand**, polymerase **epsilon** continues DNA synthesis

- In bacteria, DNA **polymerase III** is the major replicative enzyme

Major polymerase

- In human cells:
- DNA polymerase  $\alpha$  is complexed with primase initiating the synthesis of DNA, and then
- DNA polymerase  $\epsilon$  synthesizes the leading strand.
- DNA polymerase  $\delta$  synthesizes the lagging strand.

# Removal of primers

They should be removed, cuz you can't have RNA as a part of DNA



Simply..  
Polymerase I  
1) removes the primer  
2) fills in the gap left by primer removal

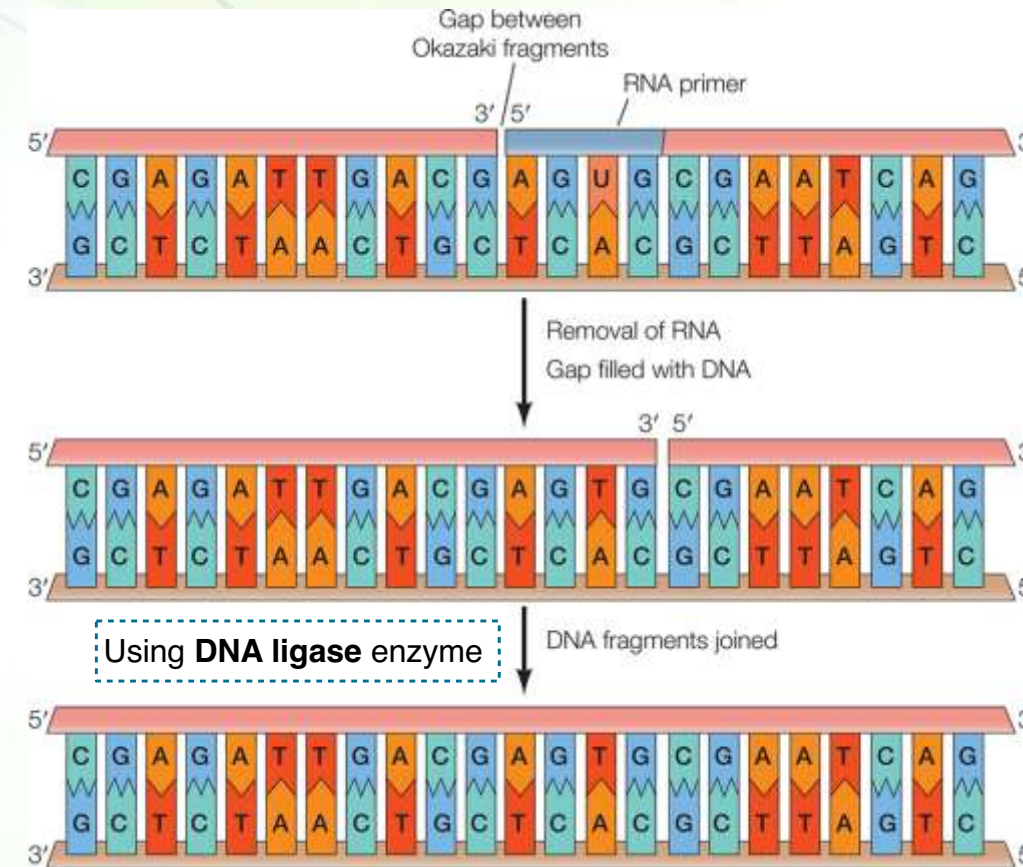
- In bacteria, RNA primers are removed by DNA polymerase I, which has two activities:

Starting removal of one ribonucleotide after another in the direction 5' → 3'

- a 5' to 3' exonuclease activity hydrolyzing the primer in the 5' to 3' direction
- A DNA polymerase activity where it fills in the gap.

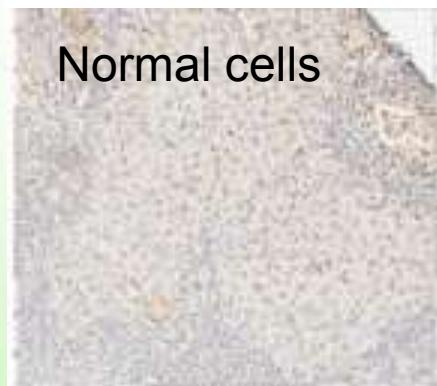
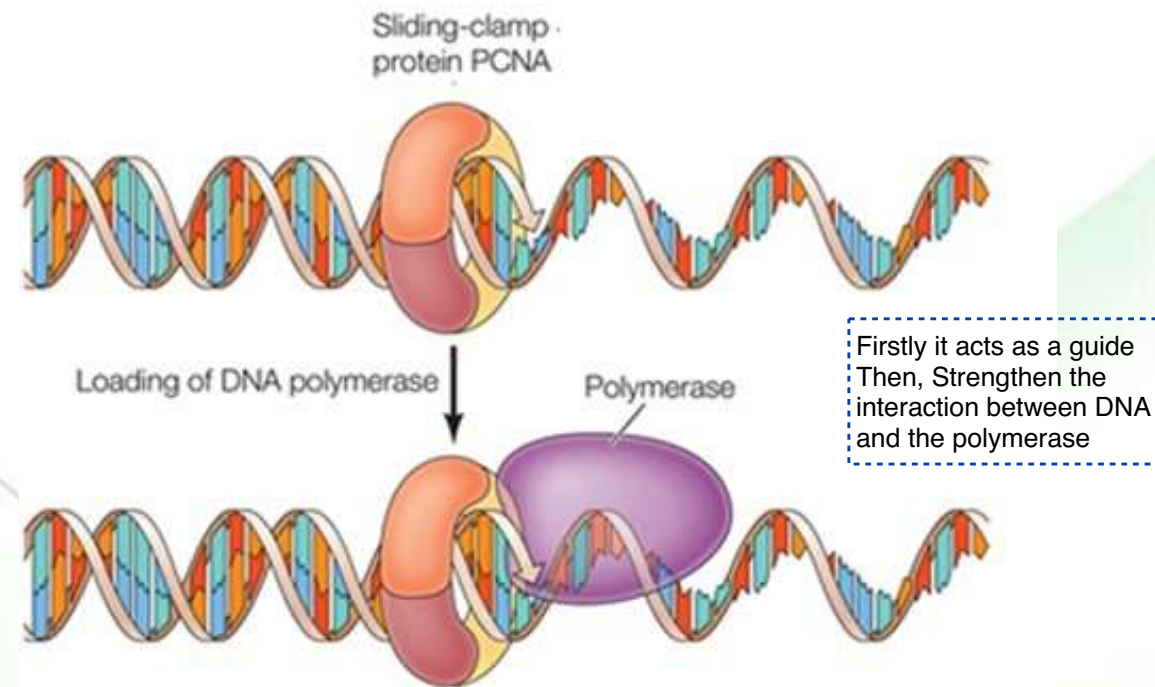
- In human cells, 3 enzymes are involved:

- RNase H**, a 5' to 3' exonuclease that removes the primers.
- polymerase  $\delta$**  that fills in the gaps
- DNA ligase** that joins the fragments.

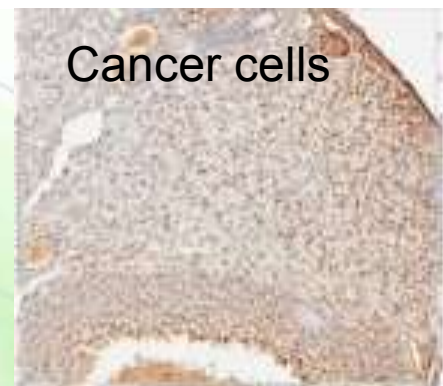


→ Removes ribonucleotides in RNA primer using **exonuclease activity**  
→ And replaces it with deoxyribonucleotide using **polymerase activity**

# Clamping and sliding



Normal cells



Cancer cells

Brown here indicates expression of proteins  
=> high expression of PCNA relative to normal cells

When Polymerases synthesise DNA they should bind to it with high stability, it shouldn't dissociate from DNA

- The sliding-clamp protein, which is called **proliferating cell nuclear antigen [PCNA]** in human cells is associated with the major replicative polymerases loading them onto the primer and stabilizing their association with the DNA template.
- Note: PCNA is a diagnostic marker of proliferating cancer cells.

PCNA is highly expressed in cancer cells

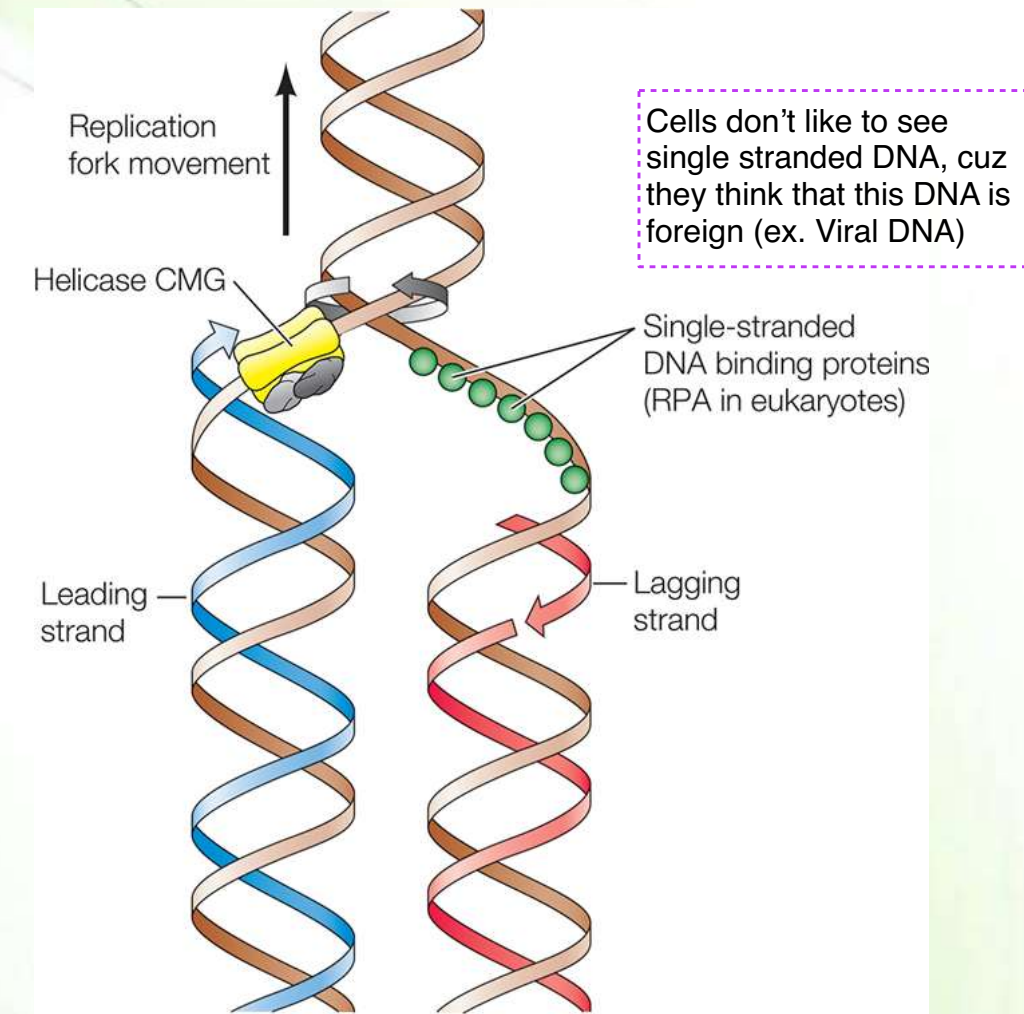
# DNA helicases and SSB proteins

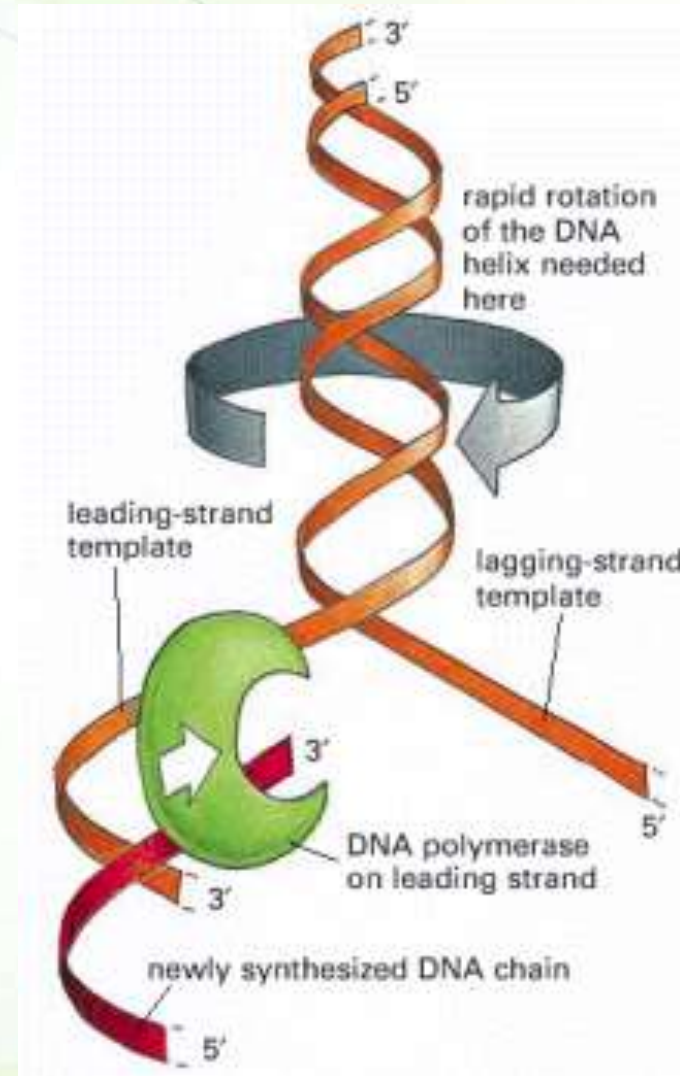
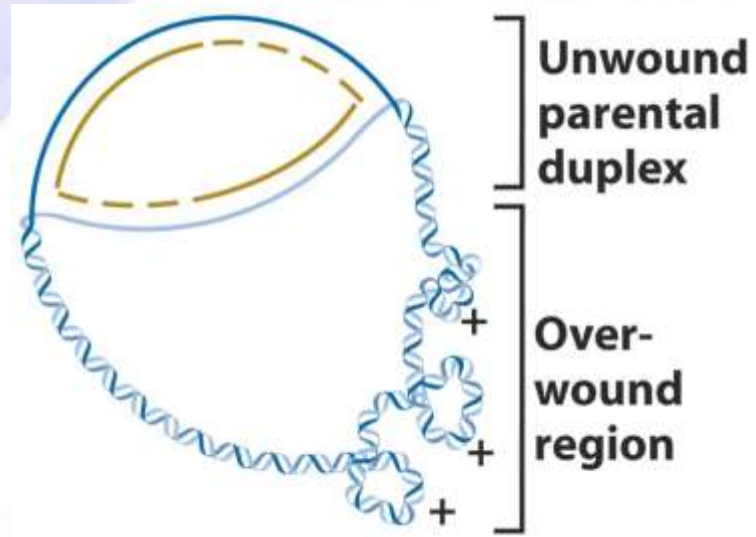


- The double-stranded DNA is opened up by **DNA helicases**. Separates the 2 strands from each other, allowing polymerase to read each one of strands

- **single-stranded DNA-binding proteins** called **replication protein A (RPA)** do these:

- 1 • Prevent the formation of short hairpin structures, 3D structure form within the same strand
- 2 • Protect single-stranded DNA from being degraded, and
- 3 • Prevent the renaturation of DNA. They could renature due to the complementary





The challenge is when DNA polymerase acts, rotation of DNA molecule occurs, either it's linear or circular DNA  
Resulting in **over-wound region (clumping of DNA)**  
Due to it, polymerases can't move forward and synthesize DNA  
So, they should be removed using topoisomerases

# DNA topoisomerases



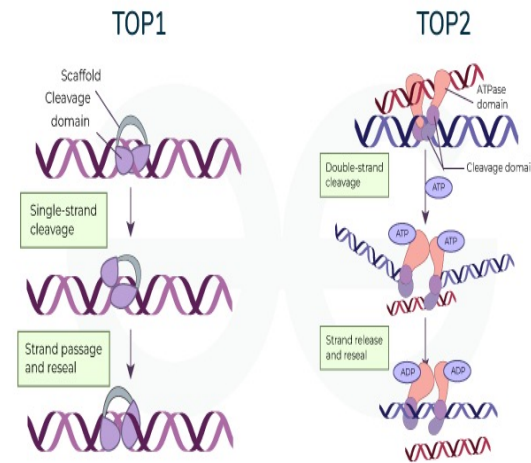
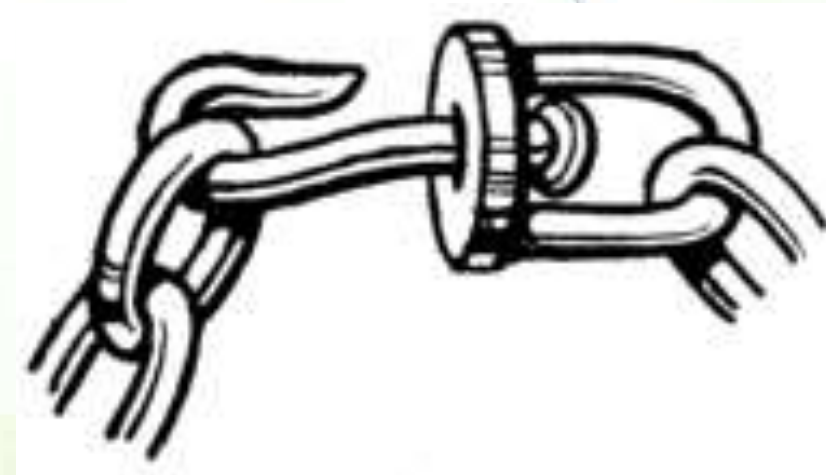
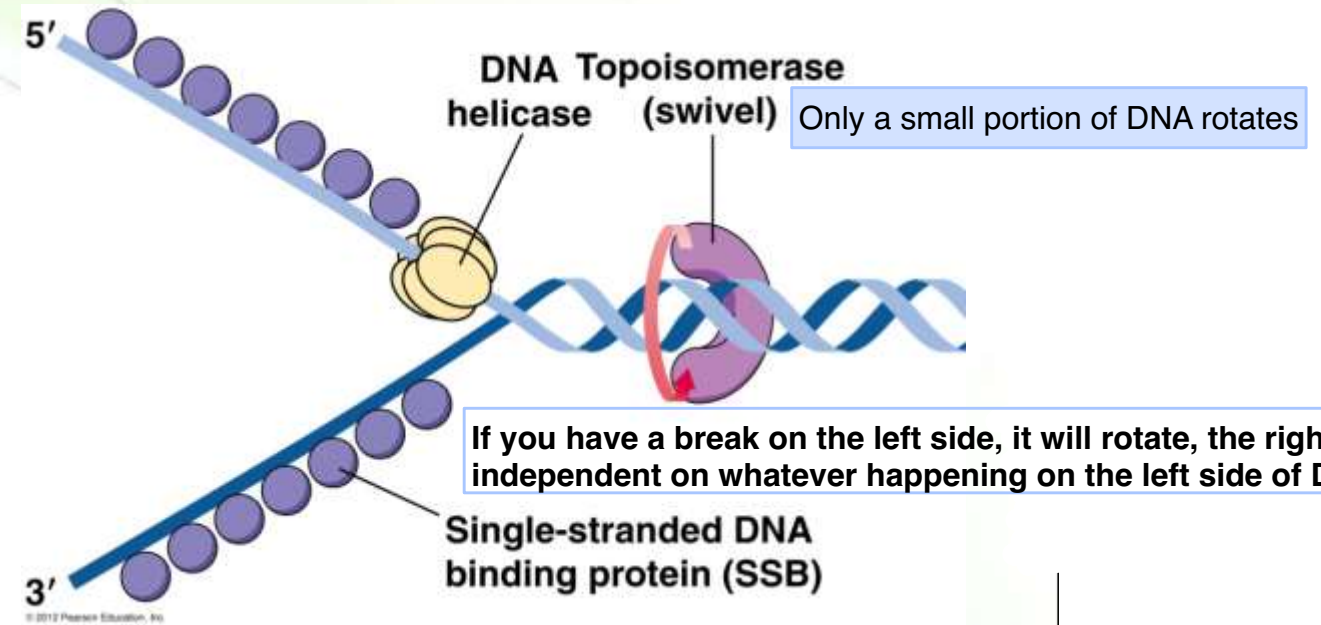
Rotation of one side without affecting the other side



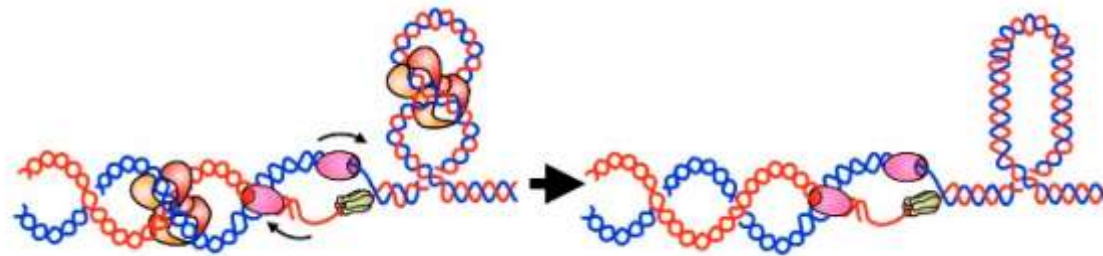
- A swivel is formed in the DNA helix by DNA topoisomerases.
- A DNA topoisomerase breaks then re-forms phosphodiester bonds in a DNA strand.

Type I topoisomerases break just one strand of DNA

Type II topoisomerases introduce two breaks: one break on each strand.



S phase  
DNA replication



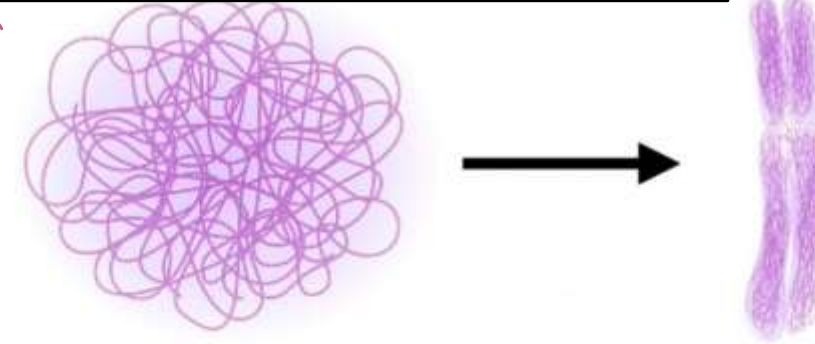


# Other functions of topoisomerase II



- Note: topoisomerase II is also required for
  - mitotic chromosome condensation
  - the separation of daughter chromatids at mitosis. Each chromatids goes in 2 opposite poles

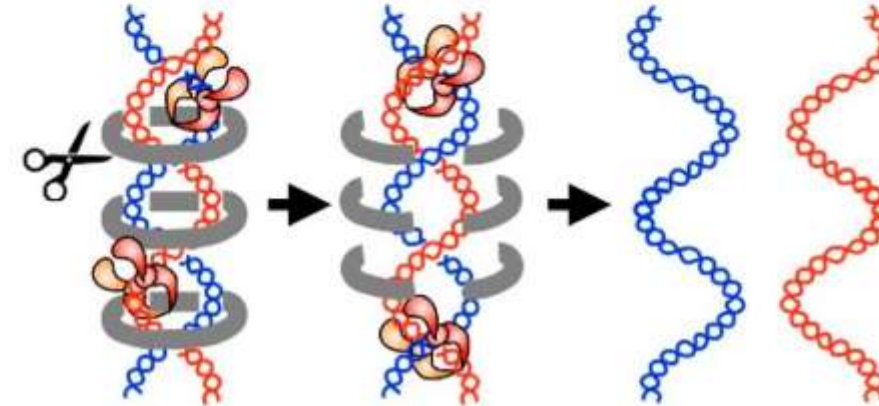
Before division of cell, converting chromatin structure into condensed chromosome (those chromosomes are separated from each other)  
This **condensation** of chromosomes is facilitated by topoisomerase II



- Antineoplastic anti-topoisomerase II inhibitors include:

- Anthracyclines
- Doxorubicin
- Mitoxantrone

All those are inhibitors of type II topoisomerase, as a result they prevent cells such as cancer cells from dividing



# How accurate is DNA replication?

- The frequency of errors during replication is only one incorrect base per  $10^9$  nucleotides incorporated.

- How is accuracy high?

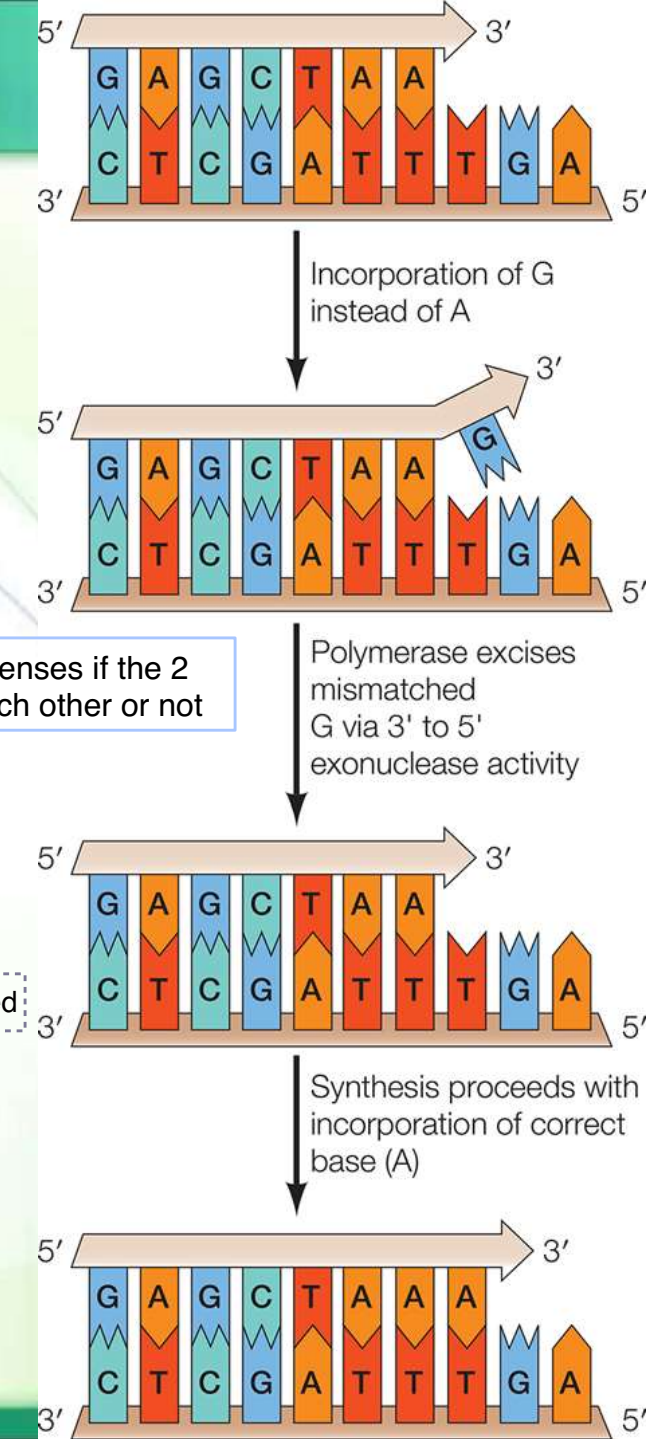
Genome consists of  $3 \times 10^9$  → means at the end of DNA replication 3 mistakes could take place in human DNA

- **The DNA polymerase** can catalyze the formation of the **right phosphodiester bonds** between the **complementary bases with the proper hydrogen bonding** (accuracy =  $10^{-5}$ ). 1 mistake could happen per 10,000 nucleotide added

The polymerase enzyme itself senses if the 2 bases are complementary to each other or not

- **Proofreading mechanism** (a  $3' \rightarrow 5'$  exonuclease activity) increasing the accuracy to  $10^{-8}$ . 1 mistake could happen per 100 million nucleotide added

- **Repair mechanisms** (to be discussed later)



Recall..  
 $5' \rightarrow 3'$  exonuclease activity, which is important for the removal of primers

# Origin of replication (OriC) in bacteria



The signal in bacteria to start replication (initiator)

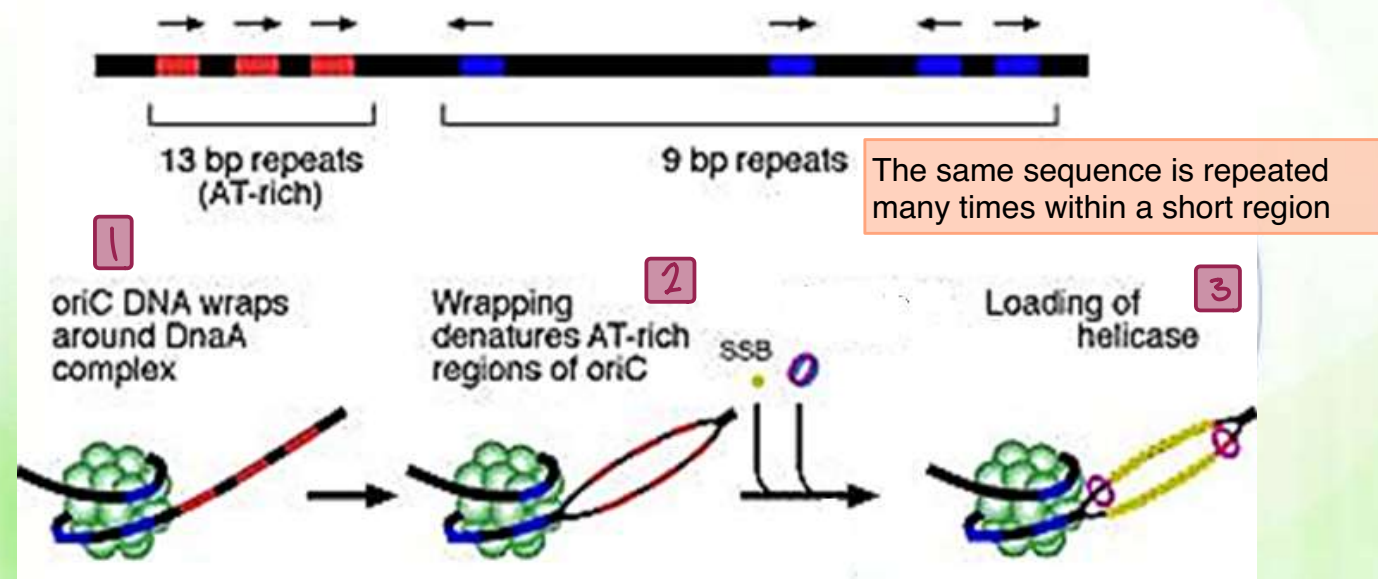
- Bacterial replication starts at an origin of replication (OriC).
- oriC regions** contain repetitive 9-bp and AT-rich 13-bp sequences (These are **consensus sequences**).

Consensus sequence:- you can find the same sequence in different genes  
**Preserved sequence**

- 9-mer**<sup>Unit</sup>: binding sites for DnaA protein.
- 13-mers**<sup>Unit</sup>: AT-rich region - it facilitates separation of the double-stranded DNA. <sup>2 hydrogen bonds</sup>

- DnaA protein binds to 9-mers, applies stress on the AT-rich region, and OriC opens up.
- The helicase and SSB proteins jump on, followed by the replication machinery.

13 mer is tandemly repeated while 9 mer not right oriC



# Origins of replication in the human genome



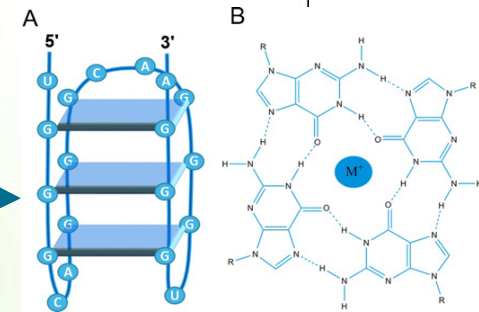
Bacteria has one site of replication OriC and it's sufficient to replicate it quite fast human genome is huge one origin of replication isn't really sufficient

They all work at the same time

- The human genome has about 30,000 origins of replication with the following characteristics:

1. G-rich sequences that can form G-quadruplex secondary structures.

4Gs forming hydrogen bonds with each other



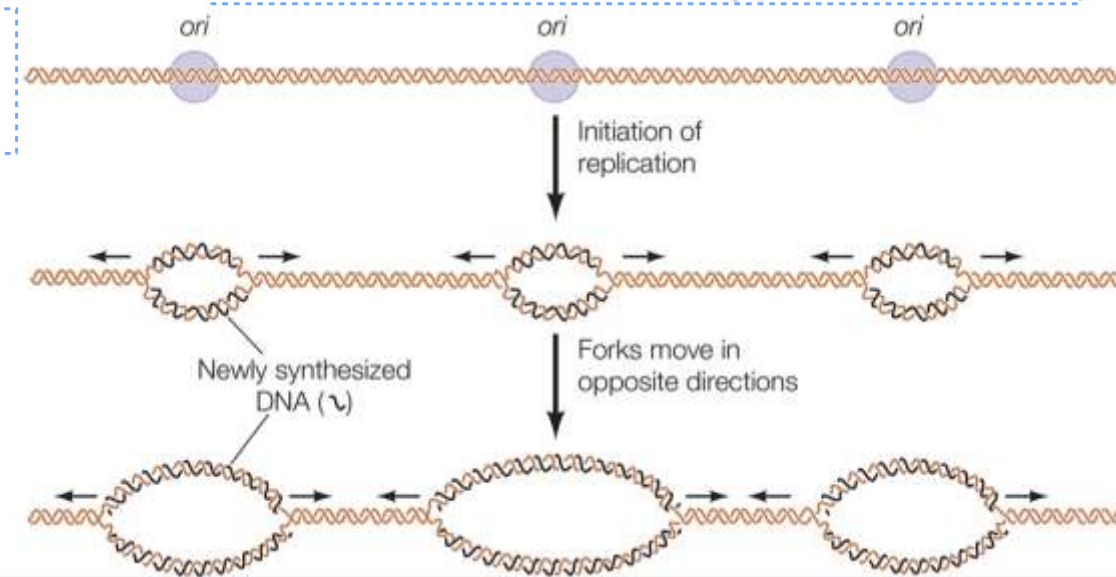
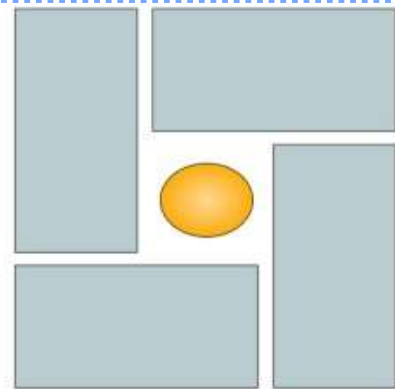
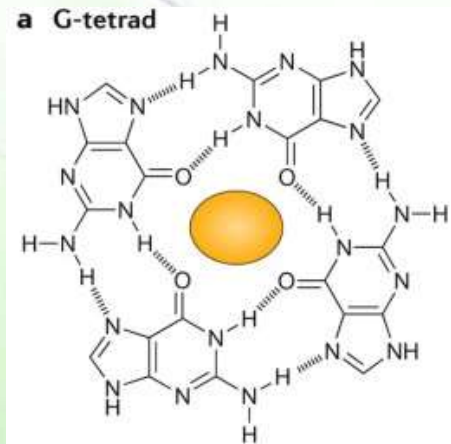
2. Modified histones that promote chromatin decondensation and activation of gene expression.

The interaction btwn histon and DNA becomes weak so that DNA can open up easily

3. Close proximity to actively transcribed genes. Origins of replication are close to actively transcribed genes

4. Cell-specific

The origins of replication in different cells are different Cuz each cell has its own active genes (specific genes) There are genes that are commonly expressed in all cells But there are genes that are cell specific



Because it replicates in a bidirectional pattern, they gonna meet together and fuse together to have newly synthesised DNA

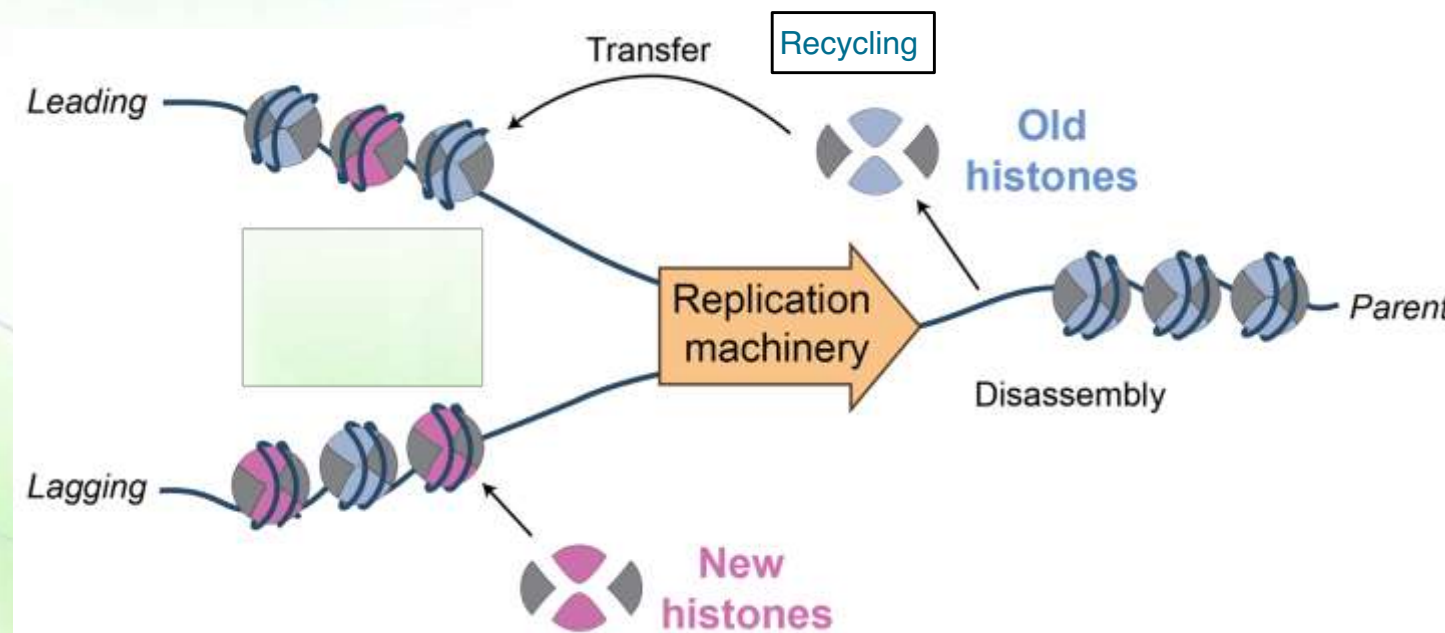
# The formation of nucleosomes



We can't have replication and transcription( to synthesize histone proteins)at the same time what should be done here?

- In the human genome, there are **4 genetic clusters** containing a total of **65 histone-coding genes**.  
*Meaning you are replicating a specific region, at the same time produce histone proteins from another region in DNA that isn't replicating*
- Nucleosomes are disassembled and reassembled during DNA replication by histone chaperones, which use recycled and newly synthesized histones.

Extra...  
organized into four clusters located on specific chromosomes:  
HIST1 cluster on chromosome 6 (the largest cluster).  
HIST2 cluster on chromosome 1.  
HIST3 cluster on chromosome 1.  
HIST4 cluster on chromosome 3.



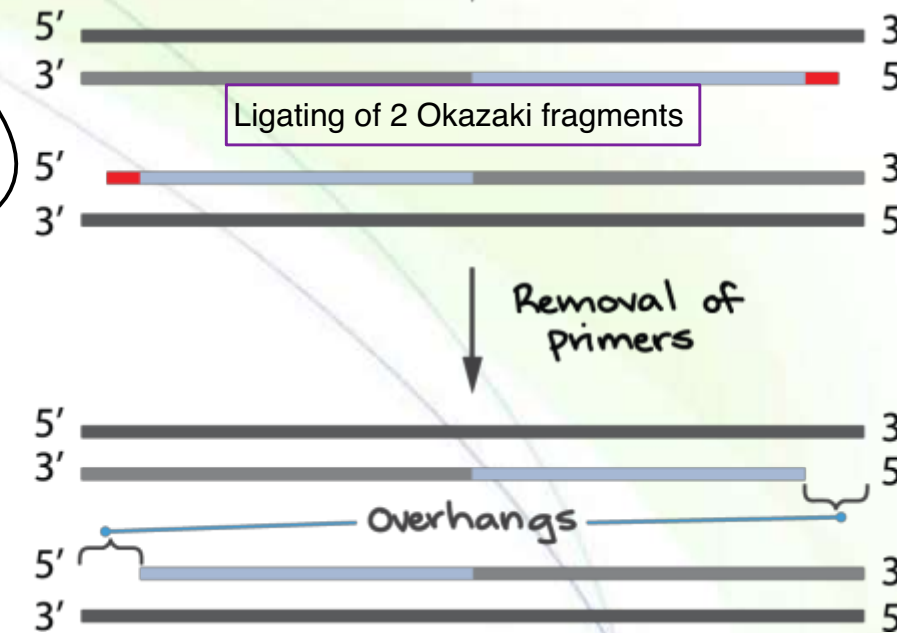
Whenever DNA is synthesised, Histone must be removed and then it must added  
As we are doubling DNA , we should also double the number of histones

# A problem in the lagging strand

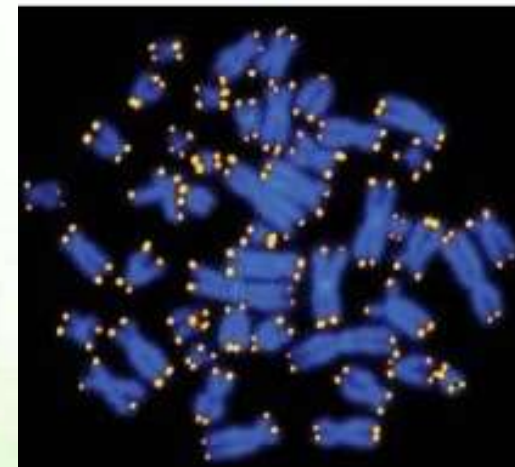


The lagging strand is not completely replicated because DNA polymerase requires a primer and cannot fill in the final gap at the chromosome end, leading to a gradual loss of DNA

- As the growing fork approaches the end of a linear chromosome, the lagging strand is not completely replicated. Why?
- When the final RNA primer is removed, there is no place onto which the DNA polymerase can fill the resulting gap leading to the shortening of the lagging strand.

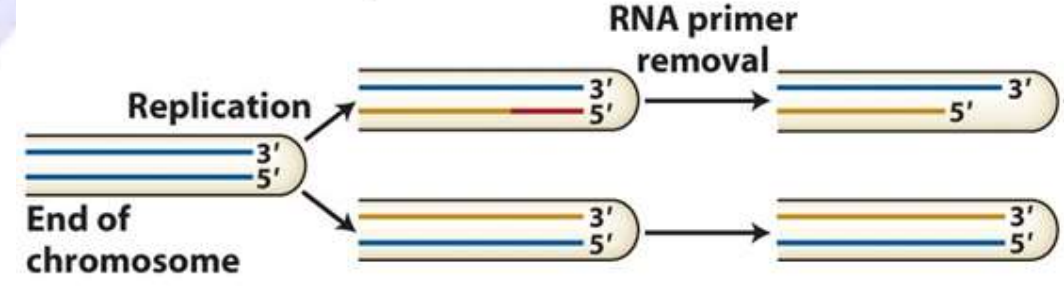


DNA polymerase doesn't have enough space to add a nucleotides, the end portion of DNA isn't replicated  
New synthesised DNA becomes shorter

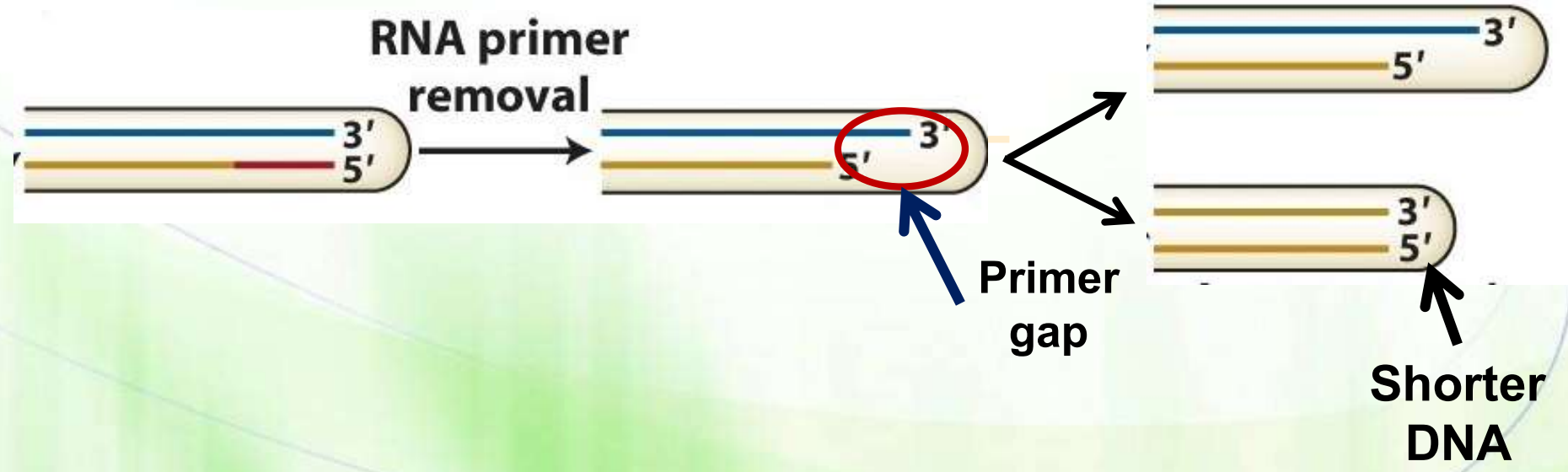
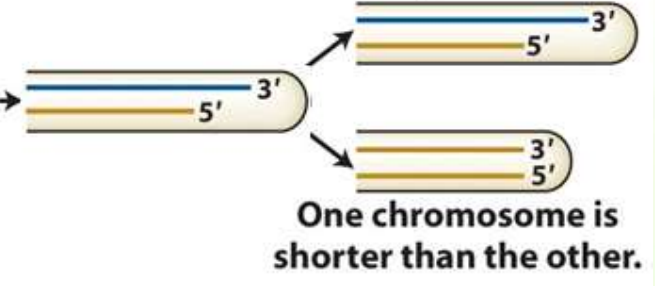




### First round of replication



### Second round of replication

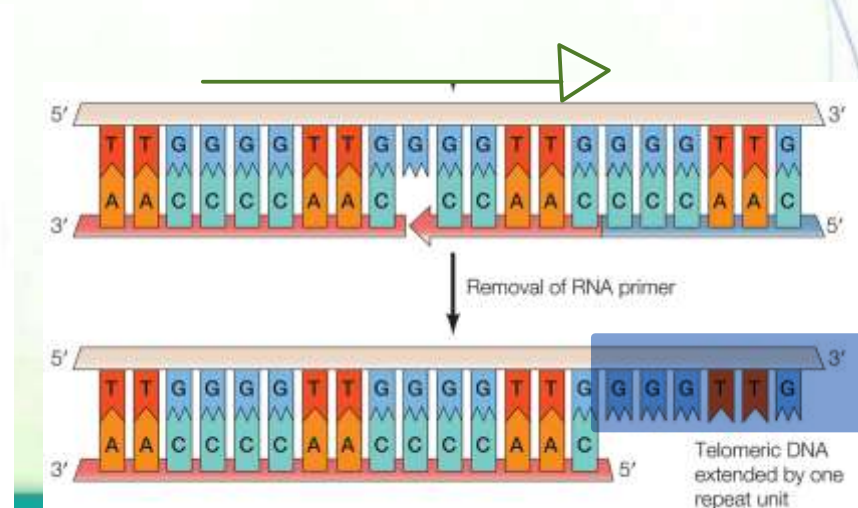
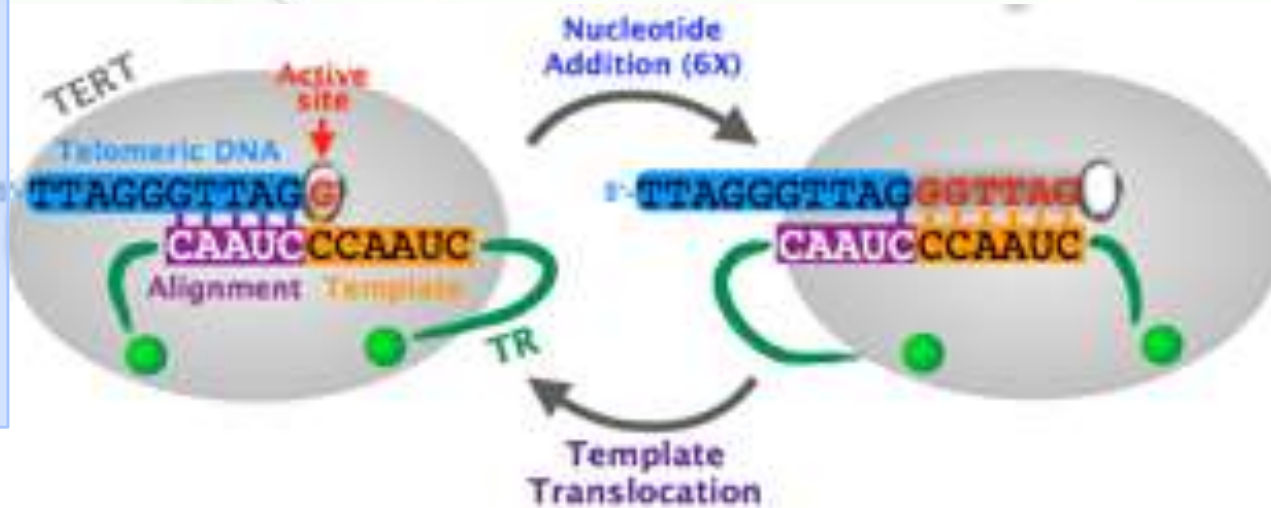


# Telomerase comes to the rescue

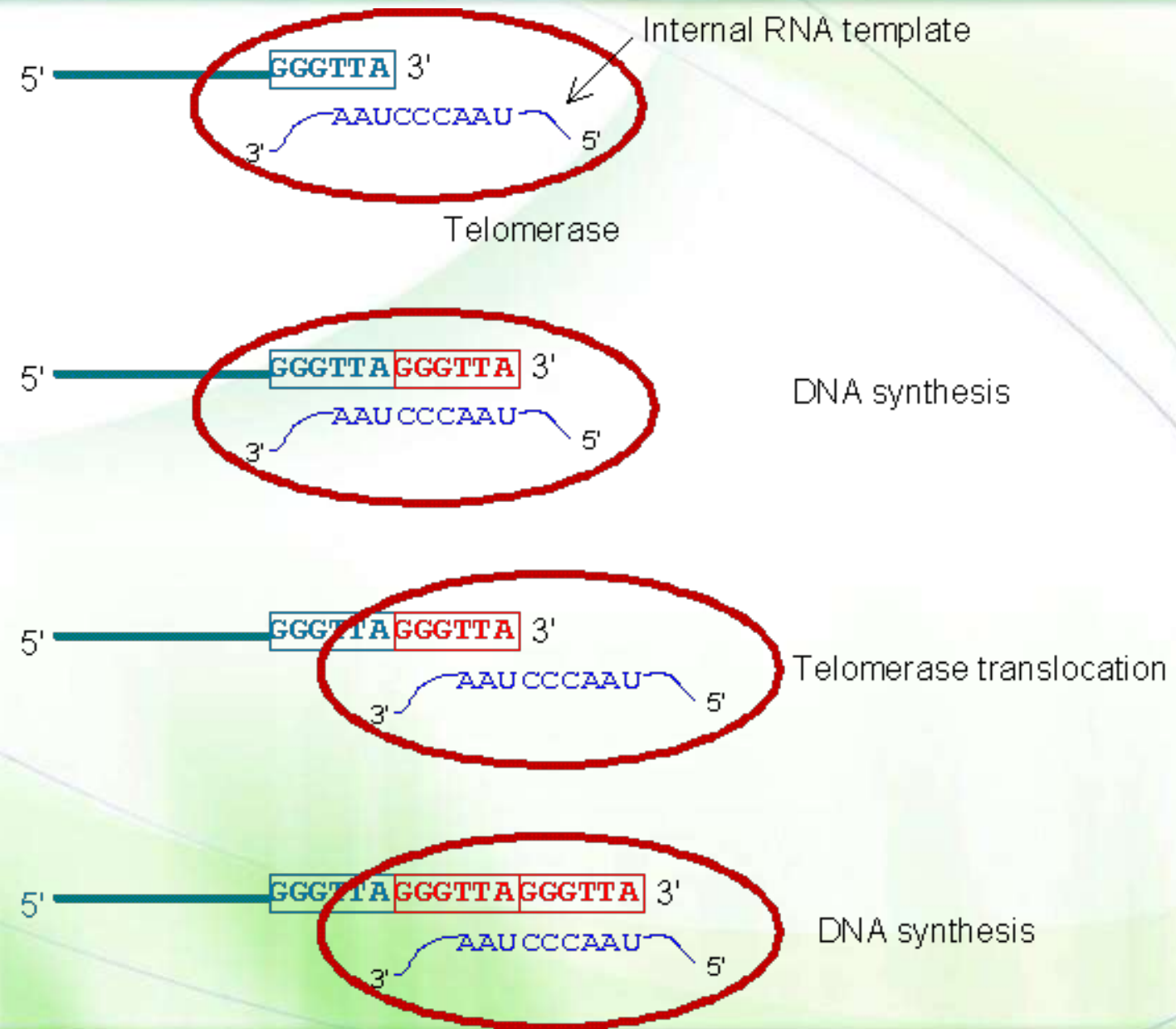


- **Elongation of lagging strand**  
Telomere DNA sequences consist of many GGGTTA repeats extending about 10,000 nucleotides. Repeats thousands to millions of times at the end of chromosomes
- Telomerase (a reverse transcriptase) prevents the progressive shortening of the lagging strand. How?
- Telomerase elongates it in the 5'-to-3' direction using a RNA template that is a component of the enzyme itself.
- When the last primer is removed, a **3'-overhang** is left.

It has its own primer within the protein itself (it's a ribonucleoprotein), it uses its primer to elongate lagging strand, so it uses RNA to synthesise DNA (because that it's a reverse transcriptase, it uses RNA template to synthesise DNA)

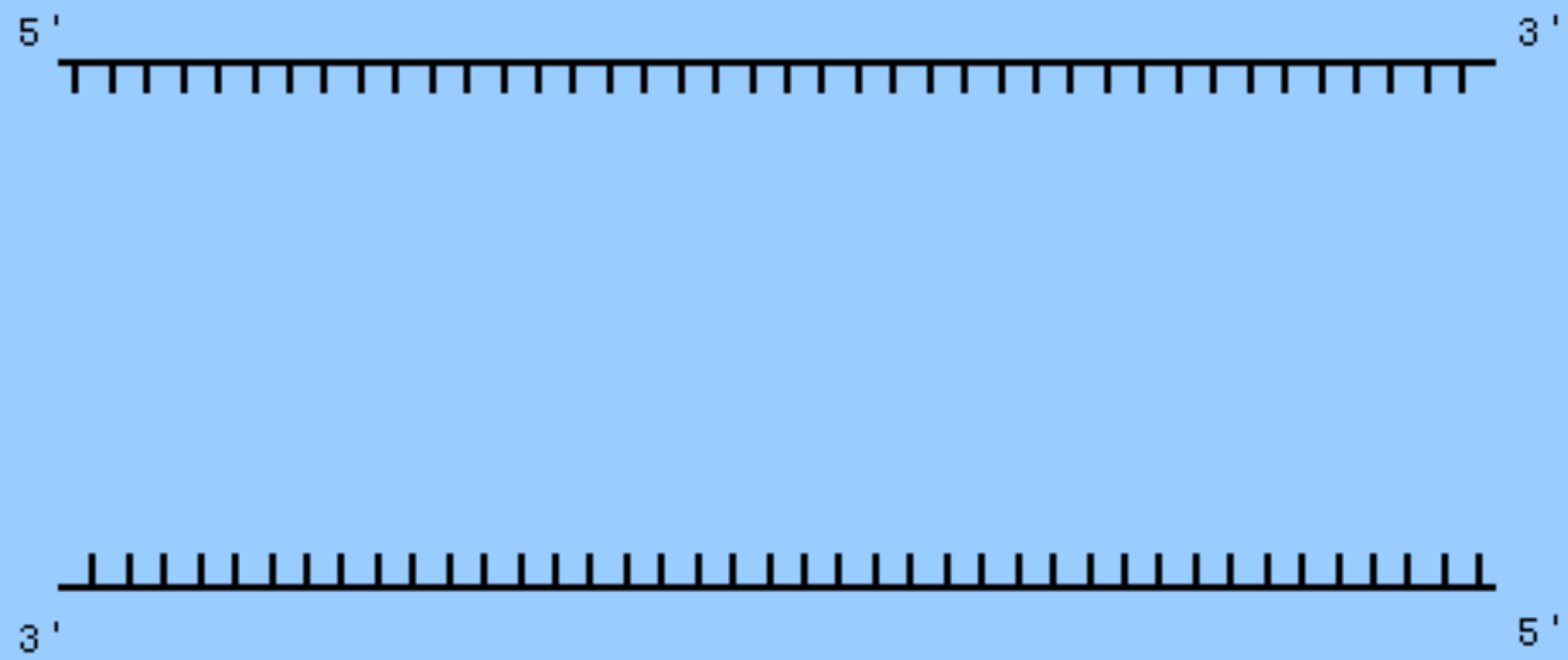








## Replication of the lagging strand of a linear chromosome encounters a problem at the 3' end

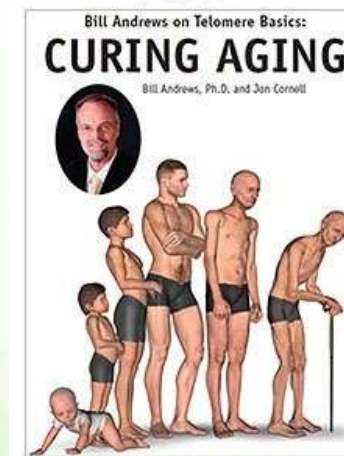
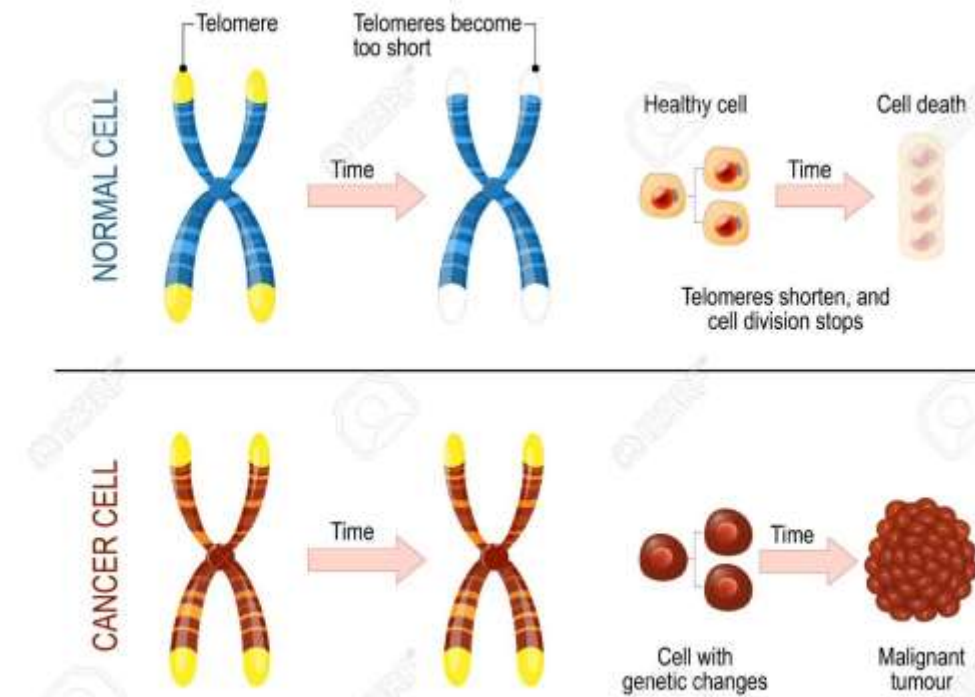


*Note: Although this animation is good, there are wrong pieces of information within it. Find them.*

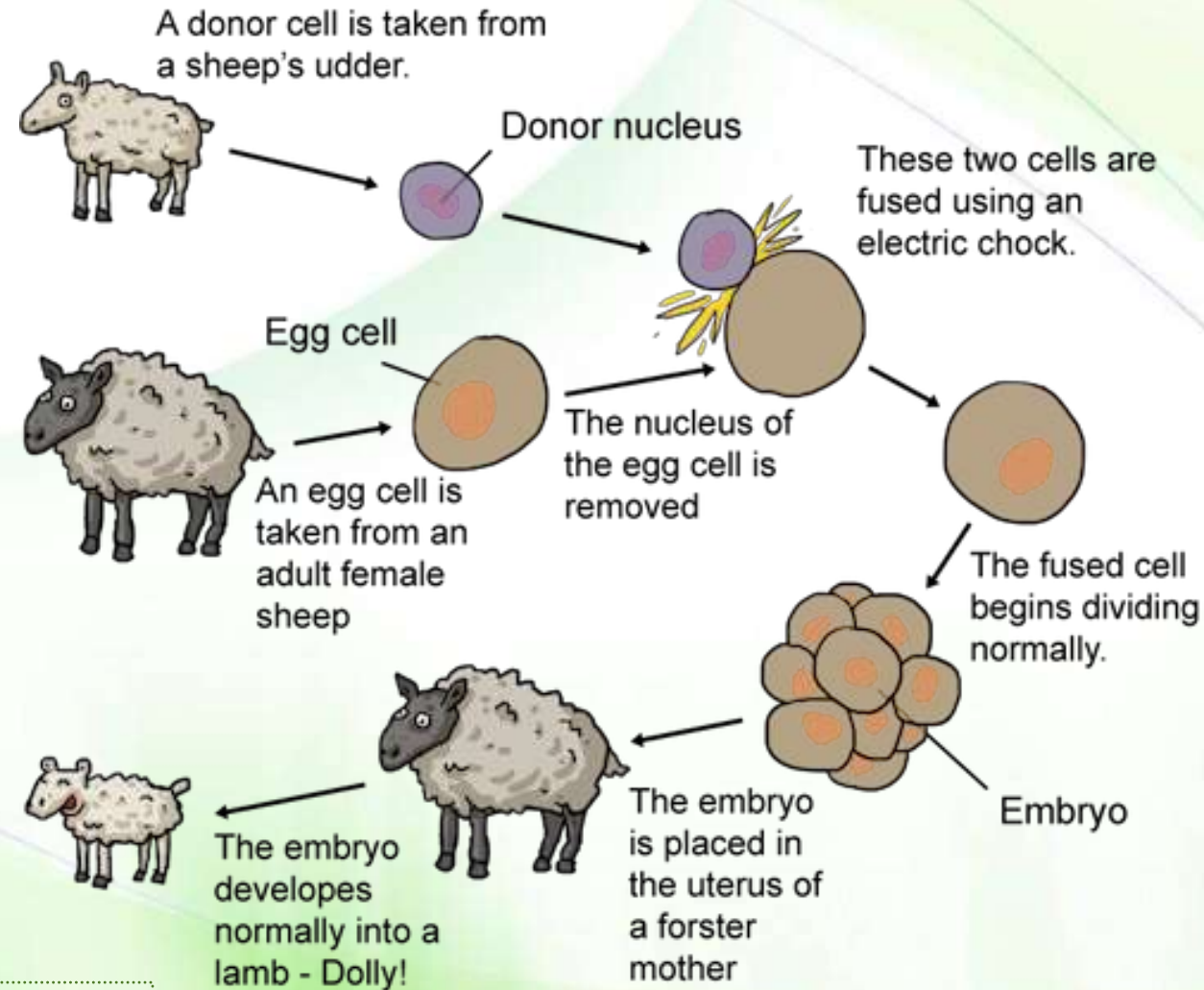
# Facts of life about telomerases



- Most somatic cells do not have high levels of telomerase and, hence, have a finite number of cell divisions.
- As we grow older, the levels and activity of telomerase are reduced.
- The gradual shortening of the chromosome ends leads to senescence and cell death.
- <sup>reproductive cells</sup> Germline and cancer cells express high levels of telomerase.



# Dolly, the sheep



Dolly lived for 6.5 years instead of the normal **11-12 years**.

Because telomeres were already 6 yrs old so new sheep lives for just 6 yrs, not like the old one which lives for 12 yrs