Restriction endonucleases, RFLP, and gene cloning

Resources

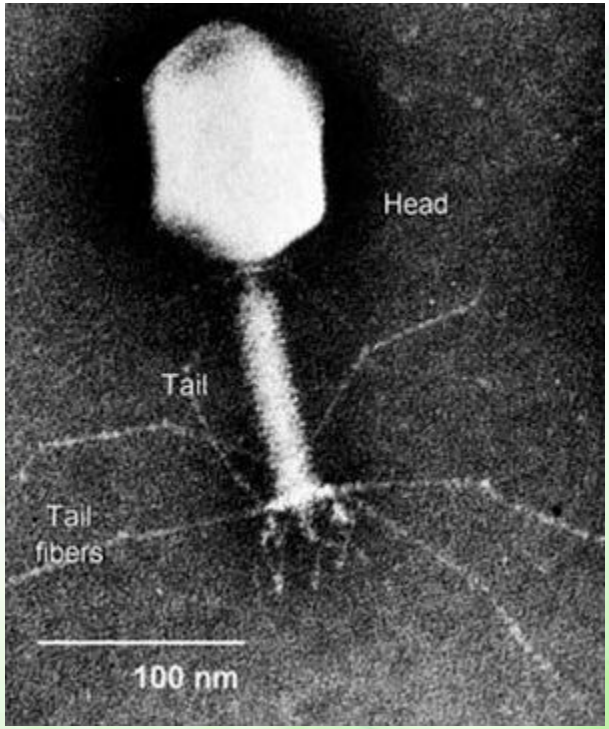
- This lecture
- Cooper, pp 120-124

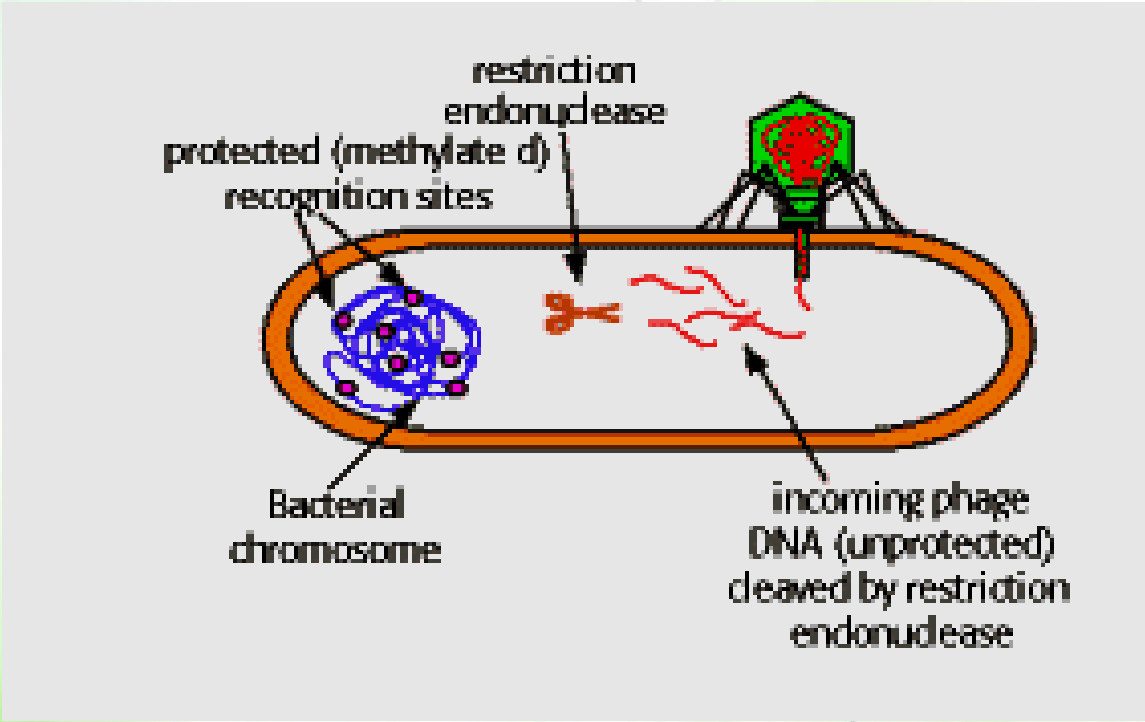
Endonucleases

- Enzymes that degrade DNA within the molecule rather than from either end (exonucleases)
- Restriction endonucleases: Enzymes that recognize and cut (break) the phosphodiester bond between nucleotides at *specific* sequences (4- to 8-bp restriction sites) generating restriction fragments.
- Type II restriction endonucleases: Always cleave always at the same place generating the same set of fragments
  - *Eco*RI (isolated from *E. coli*) cuts at 5'-GAATTC-3‘
- Some enzymes cut DNA at related sites
  - HinfI (from Haemophilus influenzae) recognizes 5'-ANTC-3' ('N' is any nucleotide)
    - Cuts at 5'-AATC-3', 5‘-ATTC-3', 5‘-AGTC-3' and 5'-ACTC-3’

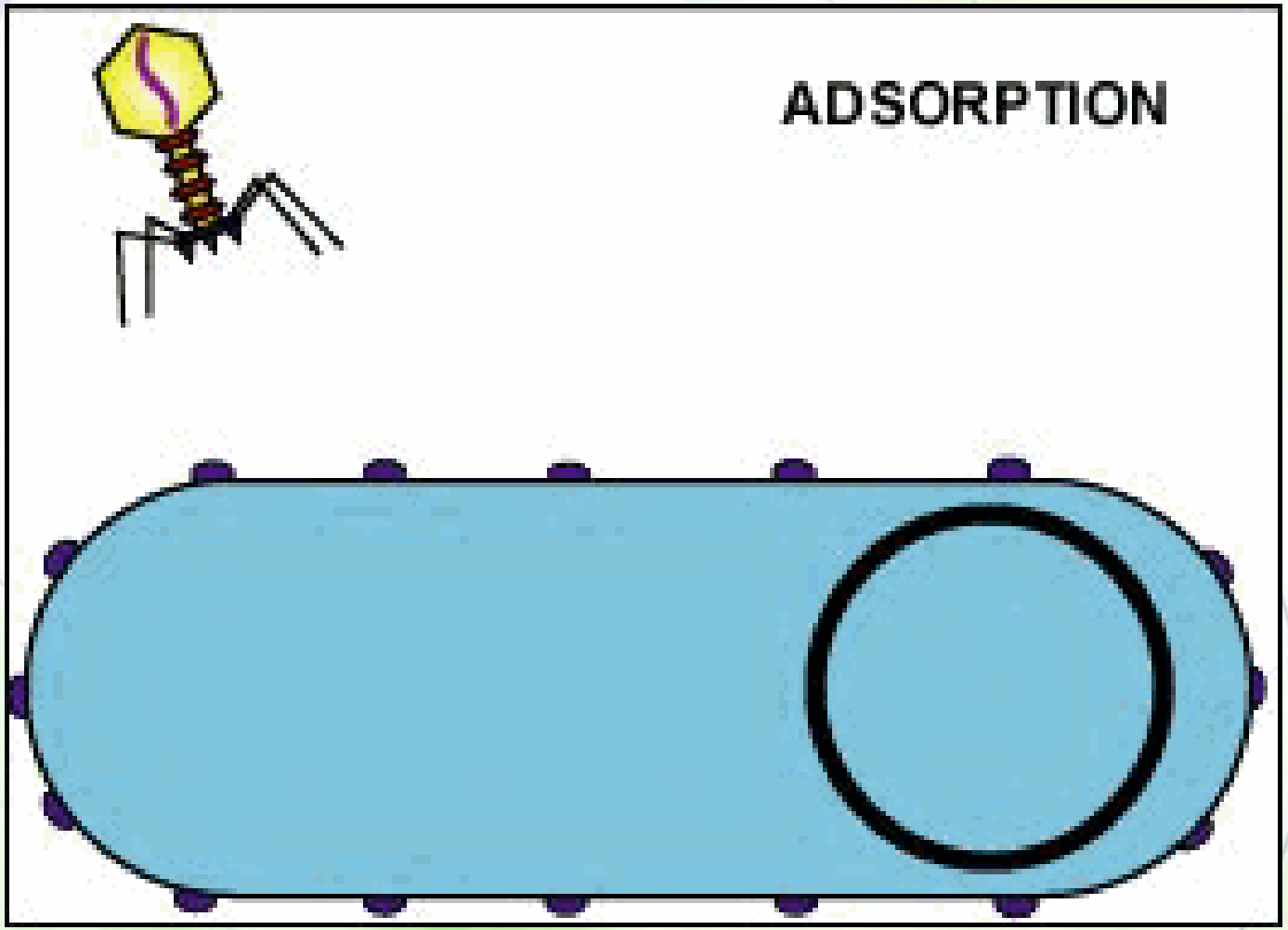
Biological purpose of restriction endonucleases

- They are present in bacteria to protect them from bacteriophages that infect bacteria by transferring their DNA into them restricting their growth.



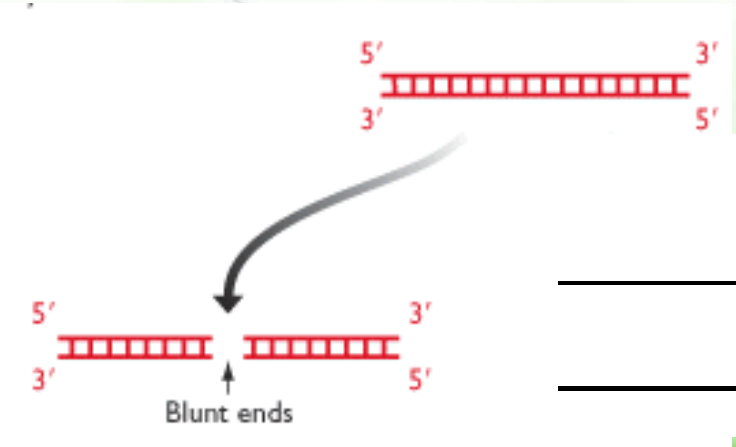


What is transduction?

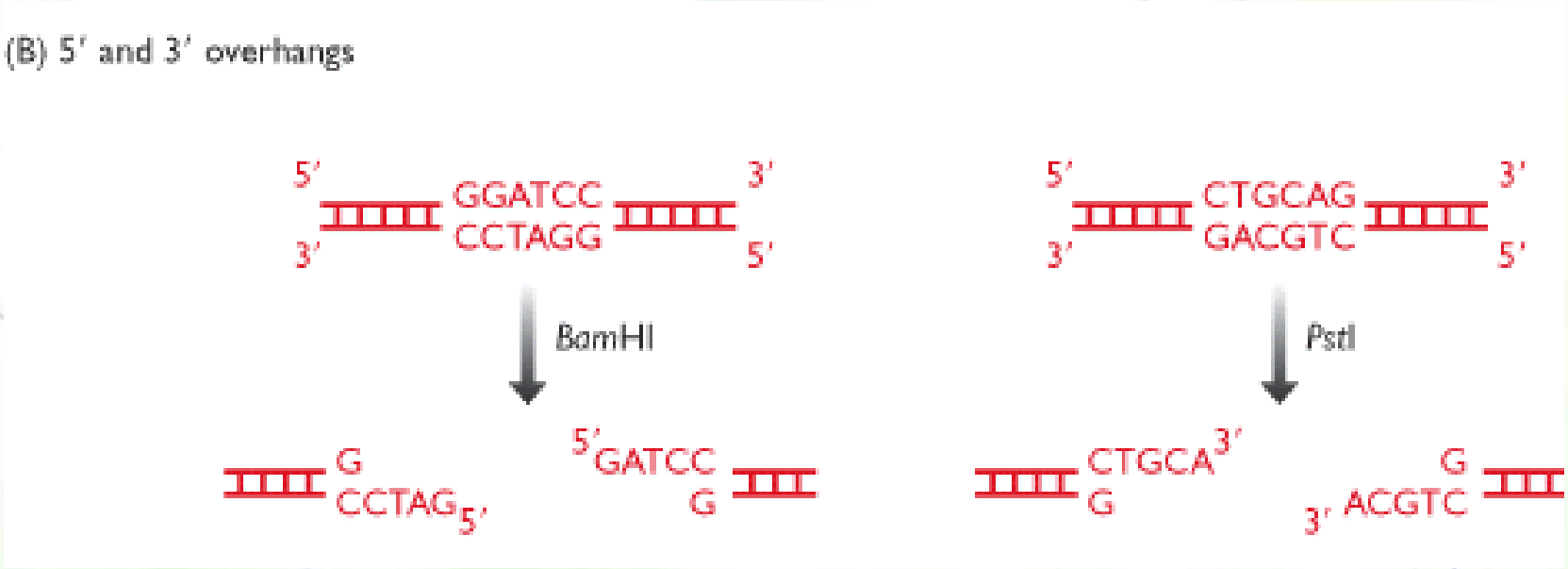


Types of cleavages

- Restriction enzymes cut DNA in two different ways:
  - Blunt: enzymes cut at the same position on both strands giving a blunt ended fragments
  - Staggered (off-center): enzymes cut the two DNA strands at different positions generating sticky or cohesive ends
    - The DNA fragments have short single-stranded overhangs at each end.



5’ vs. 3’ overhangs



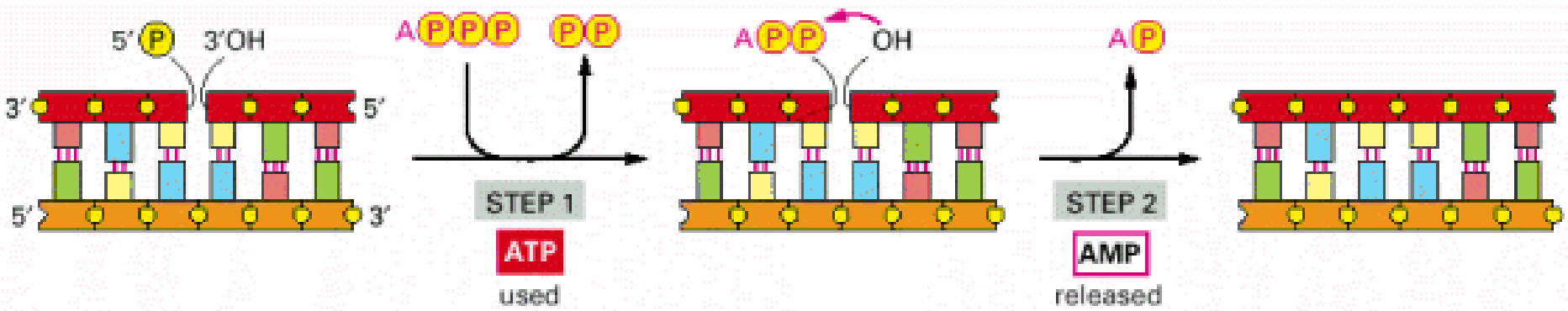
Palindromic sequences

- The sequences recognized by restriction endonucleases—their sites of action—read the same from left to right as they do from right to left (on the complementary strand).

```
EcoRI       5'  GAATTC  3'
            3'  CTTAAG  5'

HindIII     5'  AAGCTT  3'
            3'  TTCGAA  5'

SmaI        5'  CCCGGG  3'
            3'  GGGCCC  5'

TaqI        5'  TCGA  3'
            3'  AGCT  5'
```

DNA ligase

- Covalently joins DNA ends (example, restriction fragments)
- Catalyzes the formation of 3’→ 5’ phosphodiester bonds between the 3’-hydroxyl end of one strand and the 5’-phosphate end of another strand

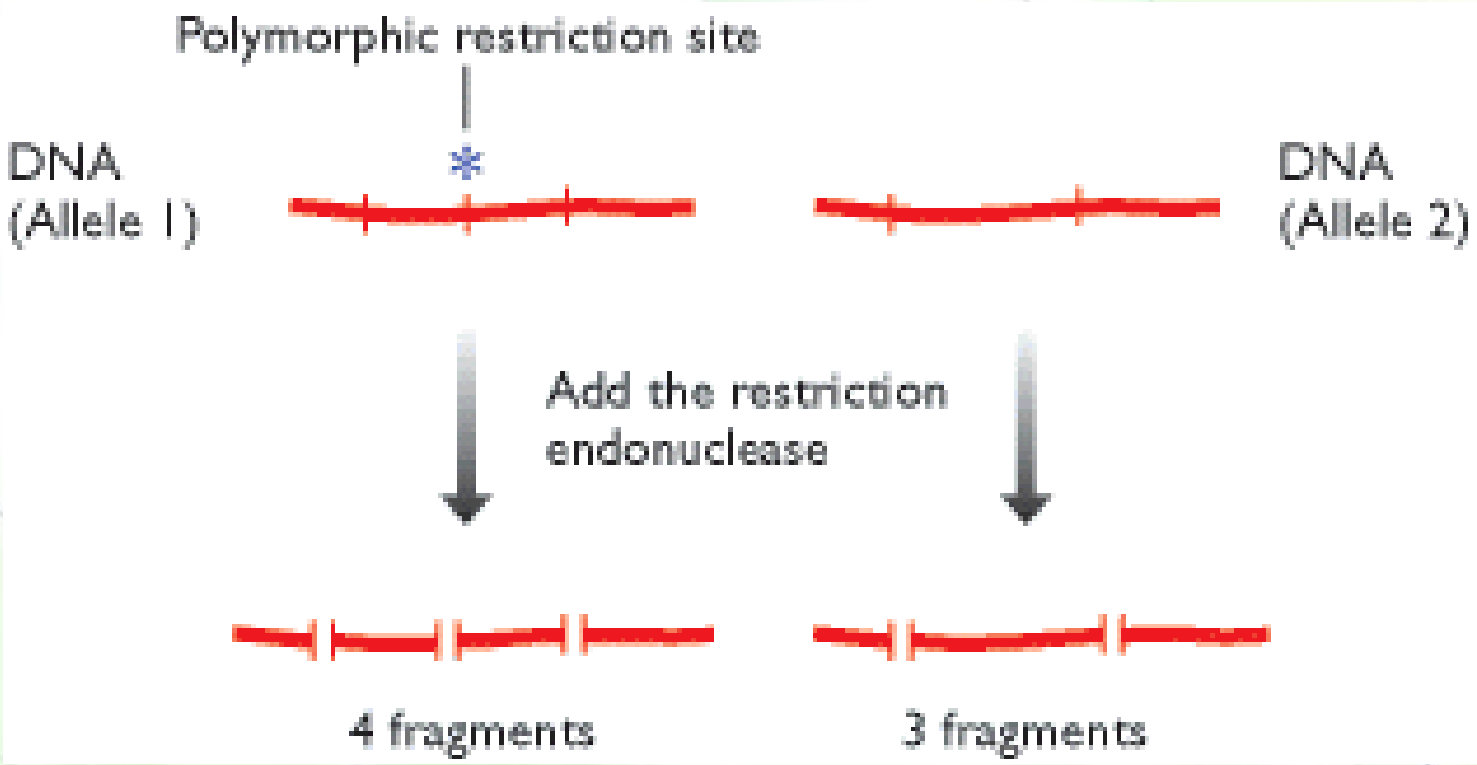


Advantage of restriction endonucleases

- Restriction fragment length polymorphism (RFLP)
- Cloning

DNA polymorphisms

- Individual variations in DNA sequence may create or remove restriction-enzyme recognition sites generating different restriction fragments
- Remember: our cells are diploid (alleles can be homozygous or heterozygous)
- What is an allele?


Polymorphic restriction site
DNA
(Allele 1)
DNA
(Allele 2)
Add the restriction
endonuclease
4 fragments
3 fragments

Restriction fragment length polymorphism

- The presence of different DNA forms in individuals generates a restriction fragment length polymorphism, or RFLP.
- These can be detected by
  - Gel electrophoresis
  - Southern blotting

Example

Variant 1

*Eco*RI does not cut

GCCGCATTCTA
CGGCGTAAGAT

Variant 2

*Eco*RI does cut

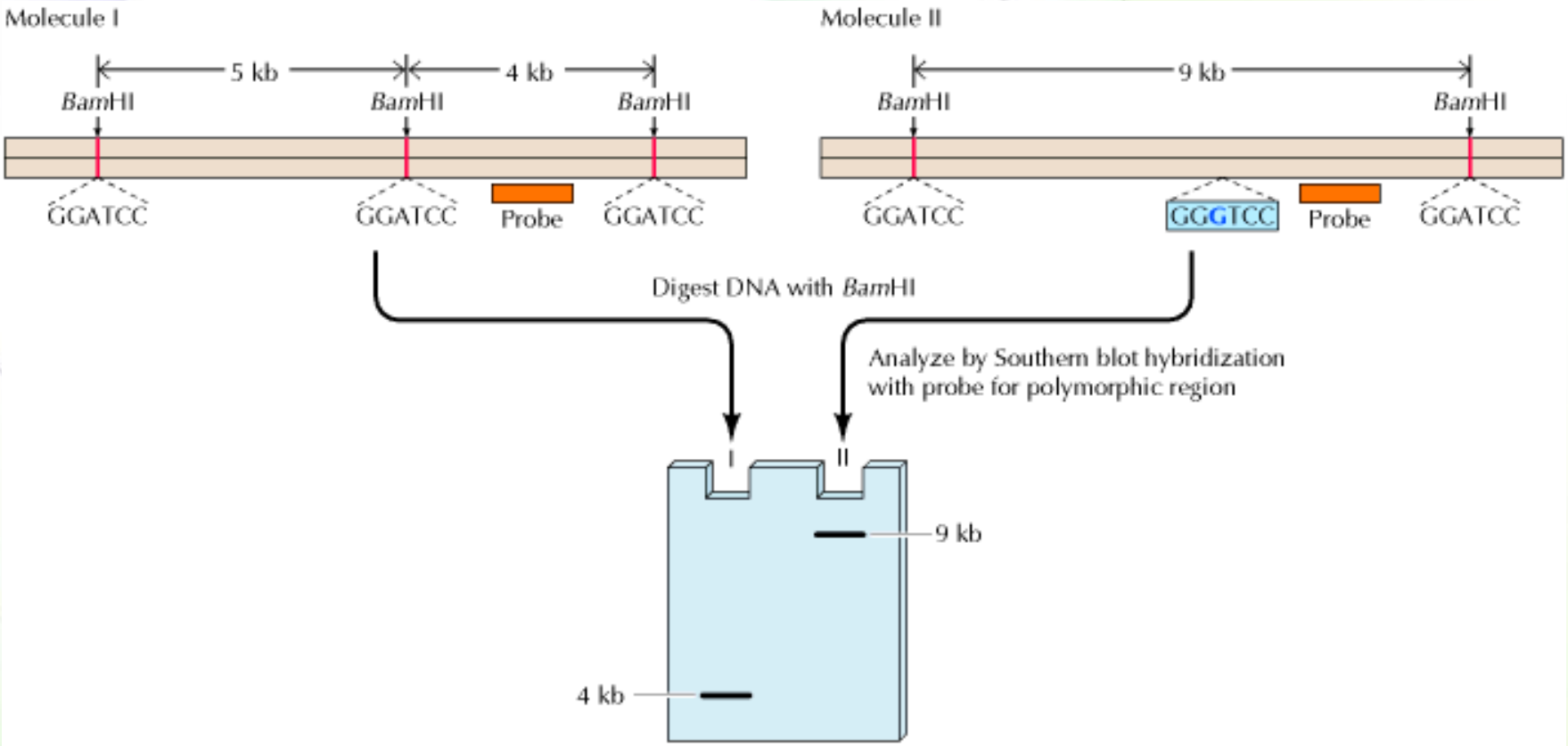
GCCGAATTCTA
CGGCTTAAGAT

RFLP in the clinic

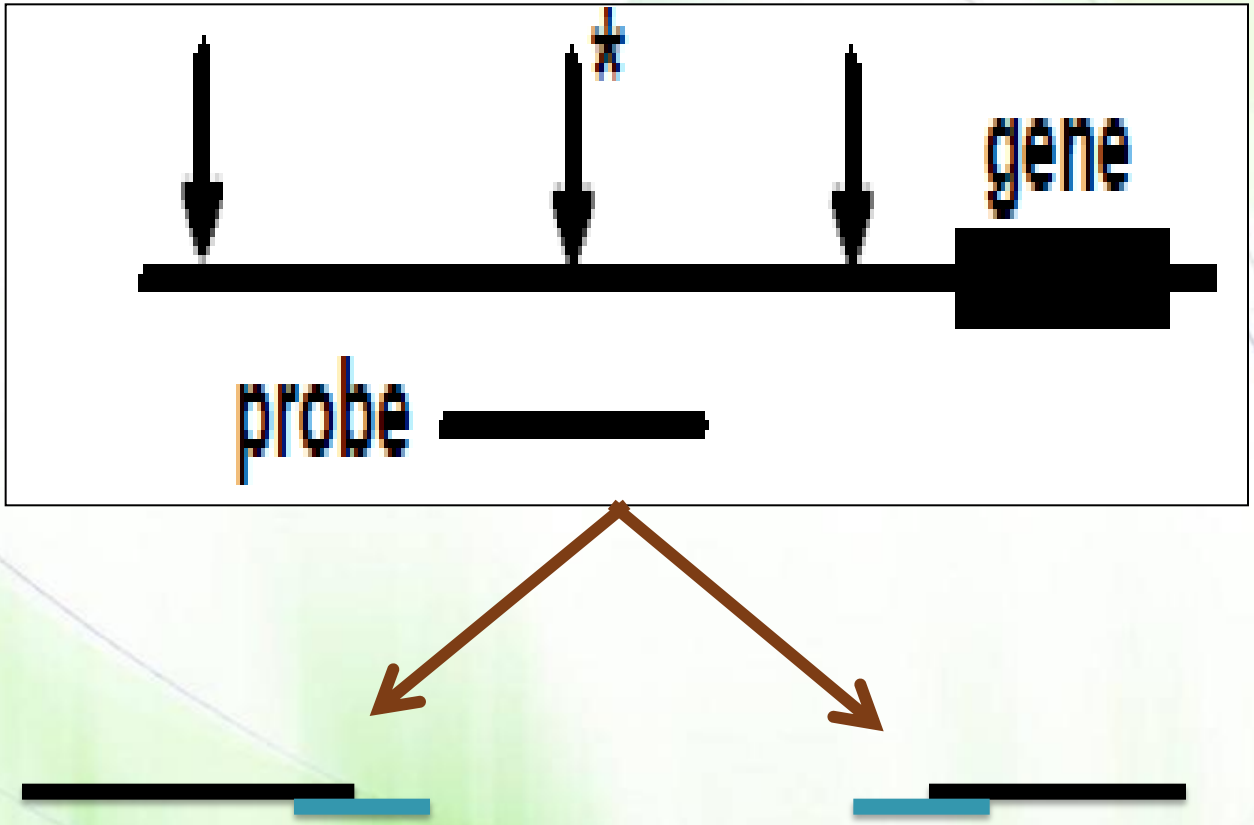
- RFLP can be used as diagnostic tools.

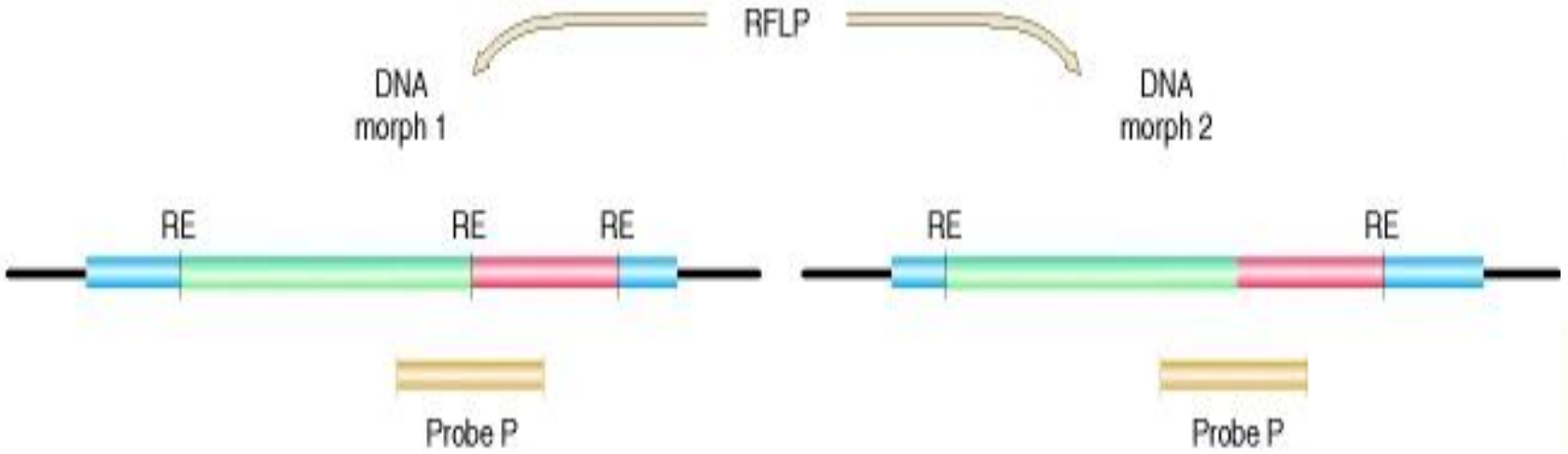
- For example, if a mutation that results in the development of a disease also causes the generation of distinctive RFLP fragments, then we can tell
  - if the person is diseased as a result of this mutation
  - from which parent this allele is inherited

Disease detection by RFLP

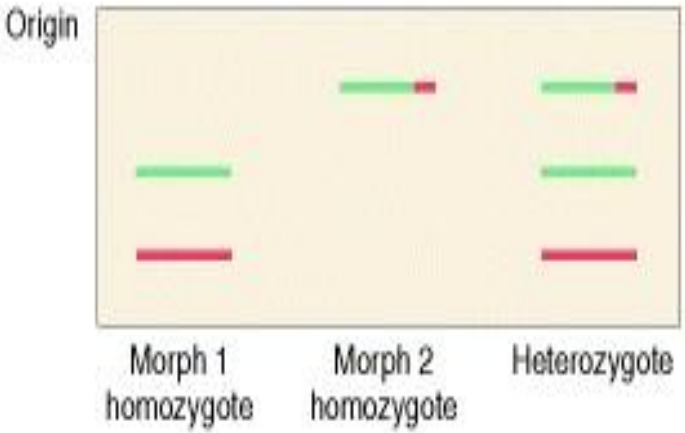


Think!! What would you see in a gel?

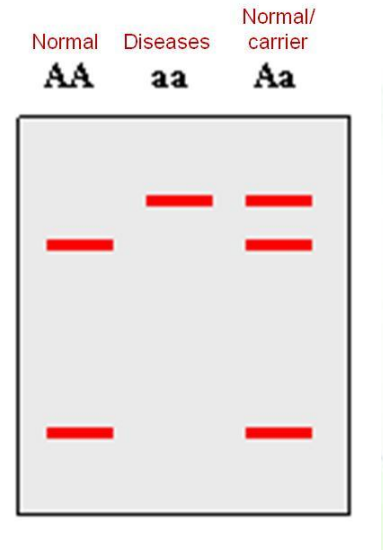


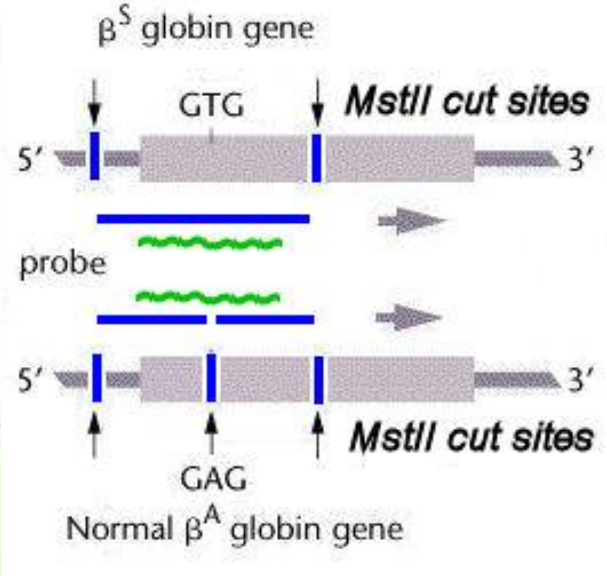

RFLP
DNA
morph 1
DNA
morph 2
RE
RE
RE
RE
RE
Probe P
Probe P

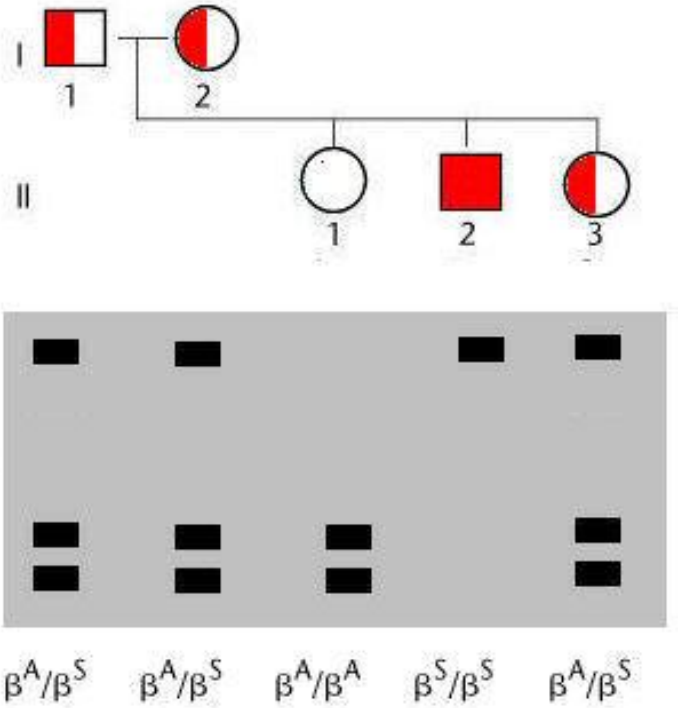
Southern


Origin
Morph 1
homozygote
Morph 2
homozygote
Heterozygote

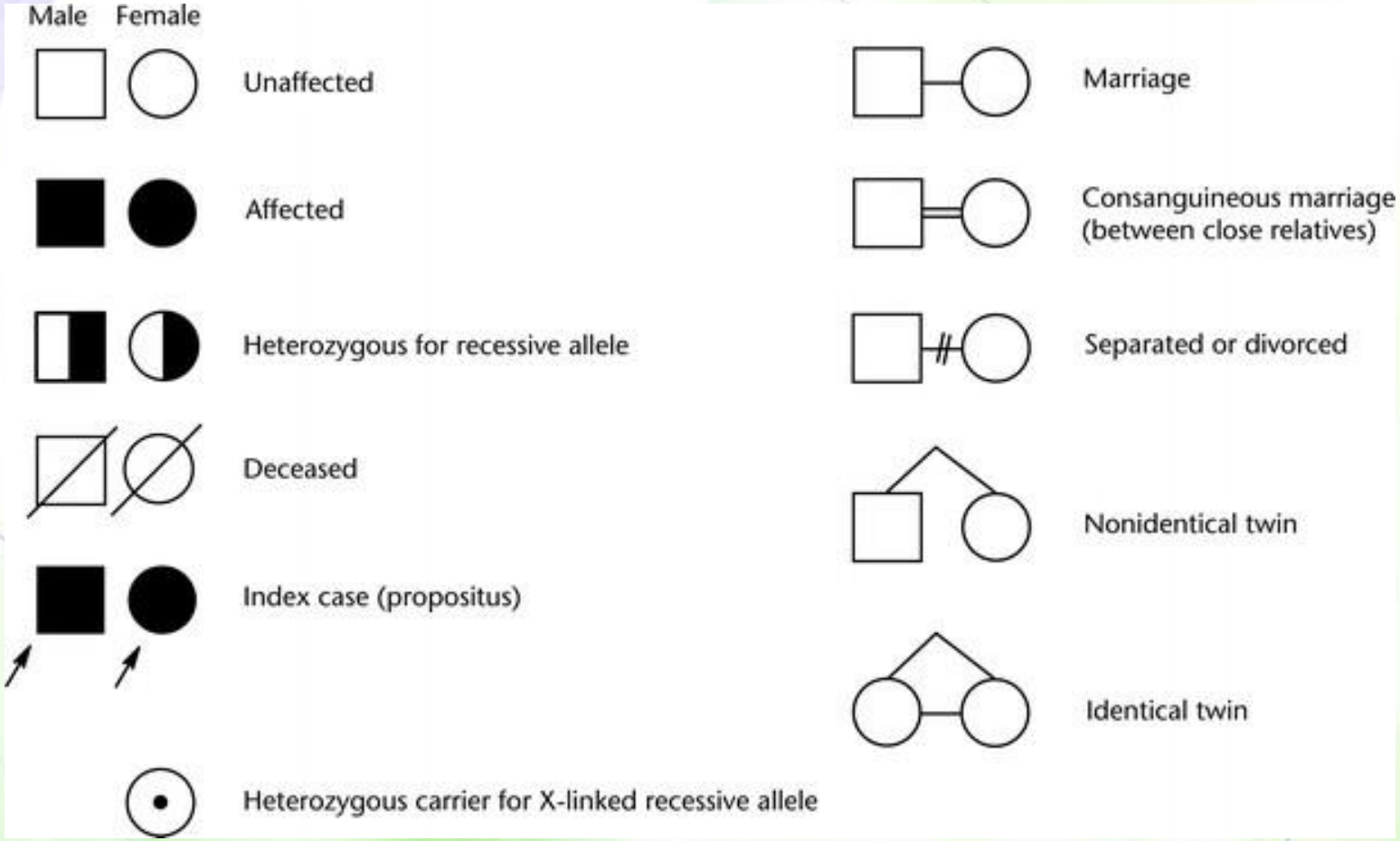
Example 1: Disease detection by RFLP (sickle cell anemia)



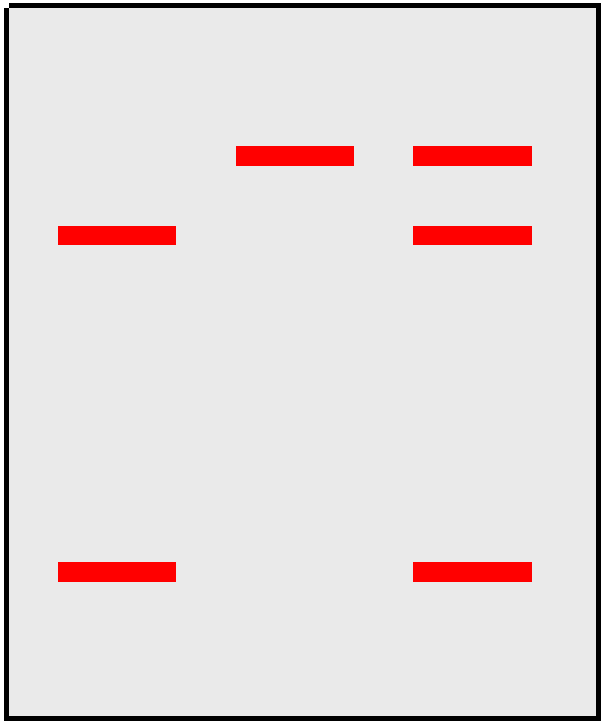




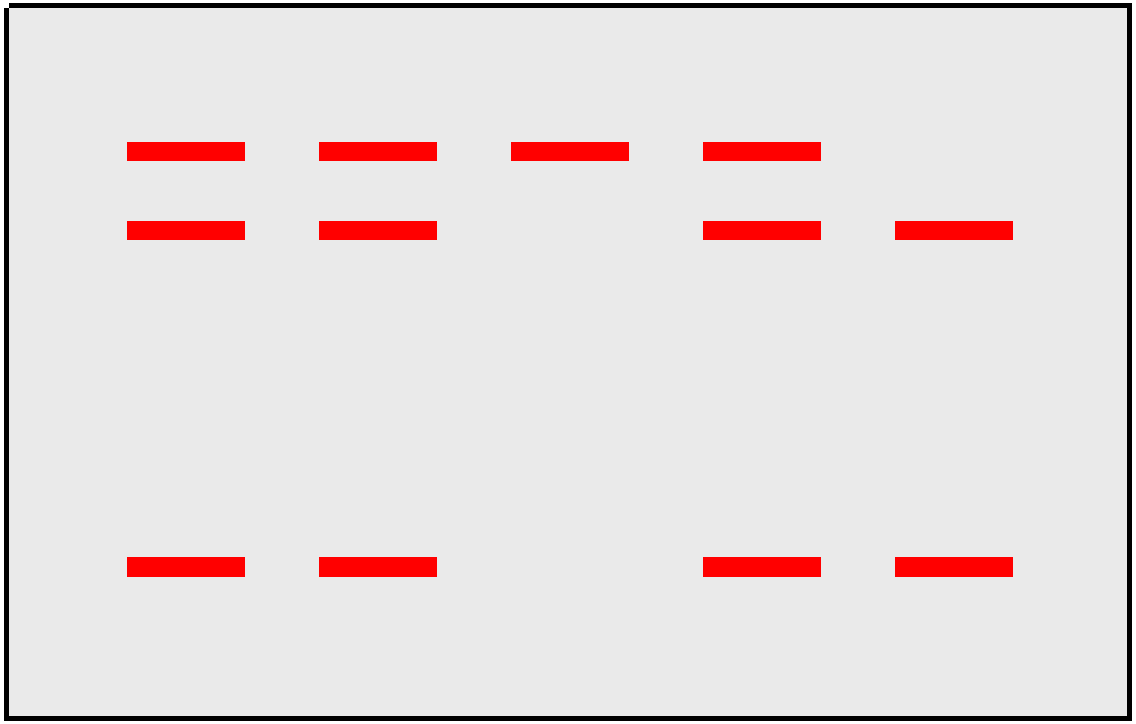
Supplementary information



Normal	Diseases	Normal/ carrier
AA	aa	Aa



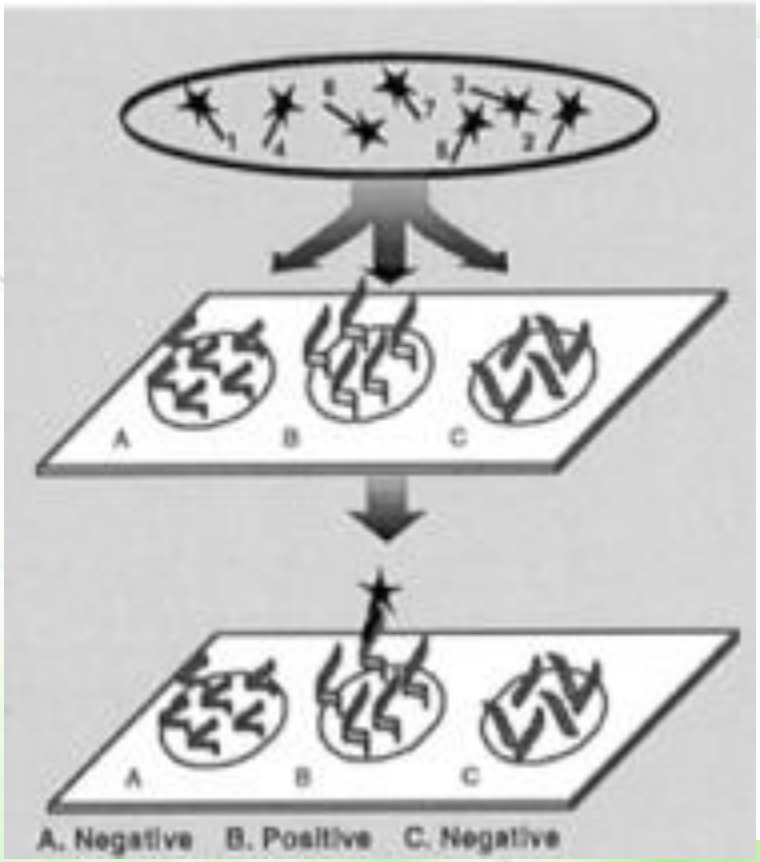
Father	Mother	Son1	Son2	Son3
Aa	Aa	aa	Aa	AA

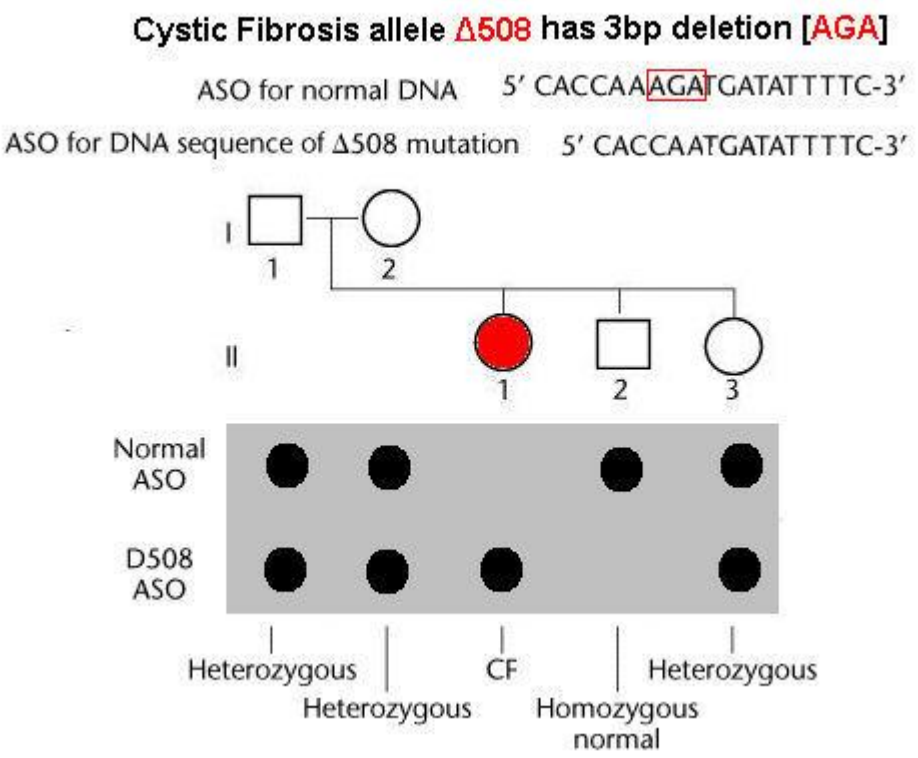


Example 2: Disease detection by ASO (Cystic fibrosis)

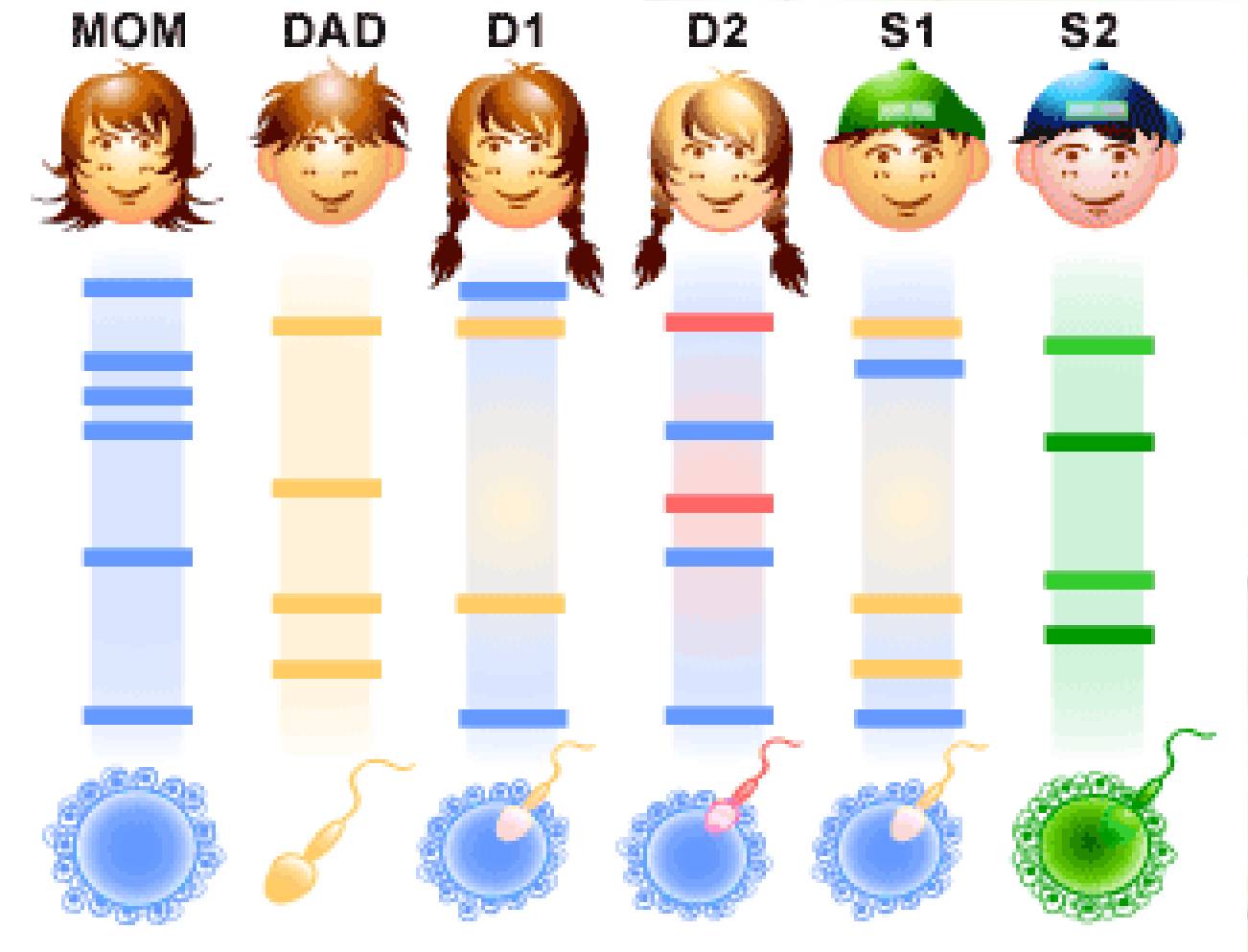
ASO: Allele-specific oligonucleotide

No electrophoresis is done here, but the whole DNA is spotted on a membrane and hybridized with two ASO's, one at a time.

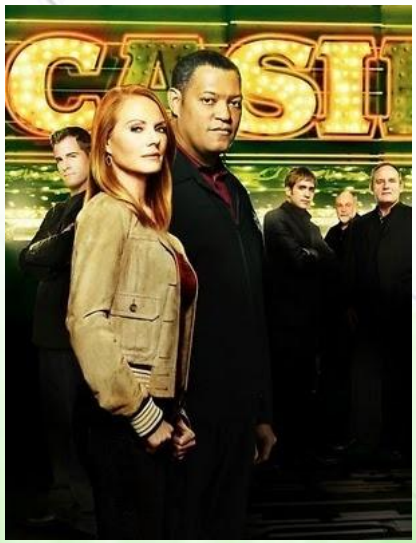


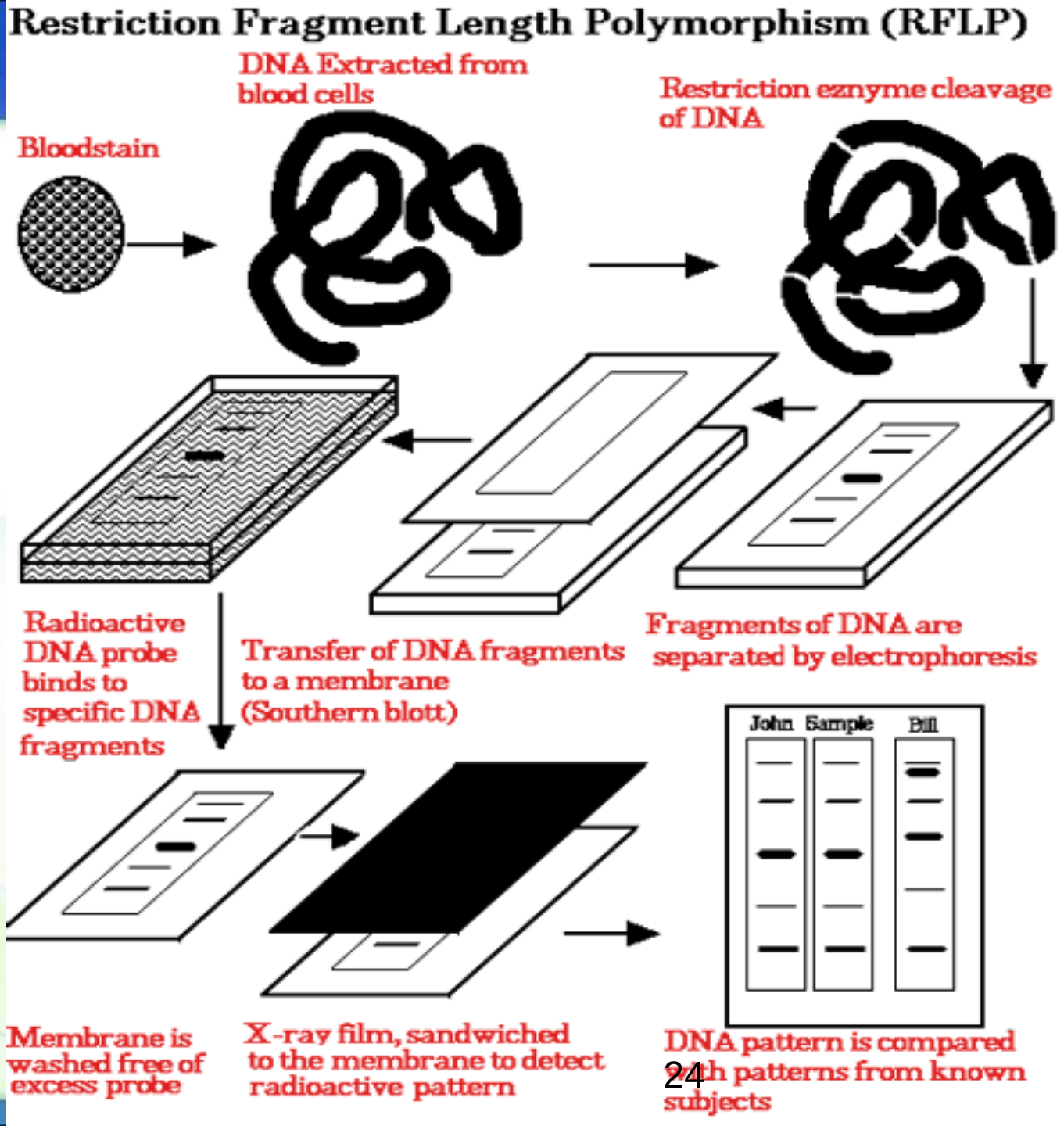


Example 3: Paternity testing



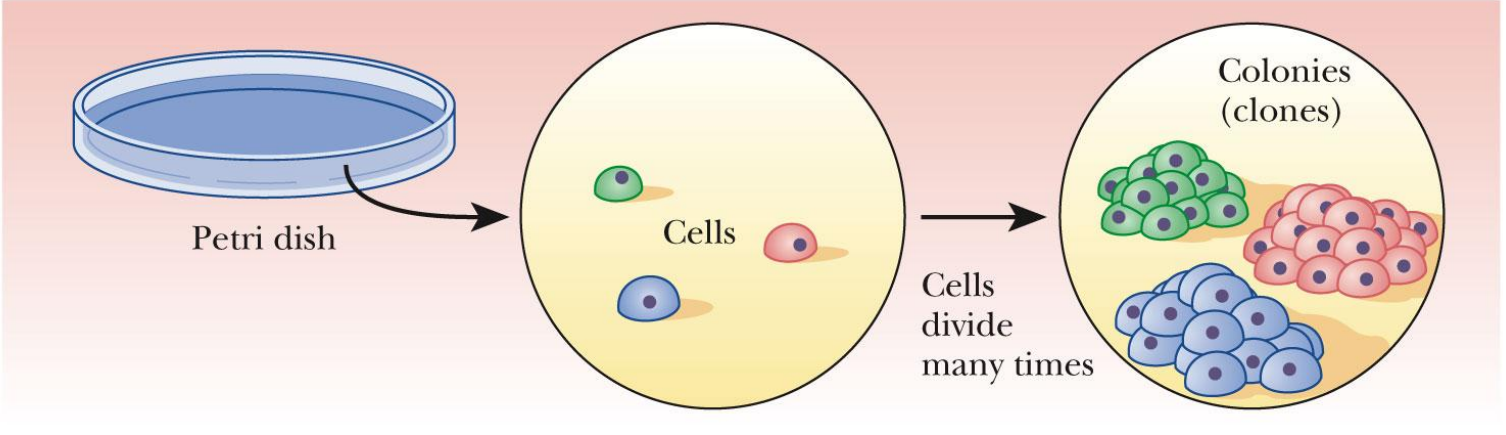
Example 4: Forensics





Cloning

- Cloning means that you make several copies of one thing.
- A clone is a genetically identical population, whether of organisms, cells, viruses, or DNA molecules.
- Every member of the population is derived from a single cell, virus, or DNA molecule.



How do we clone a DNA molecule?

- a DNA fragment of interest is inserted into a DNA carrier (called a vector) that can be replicated.
- The resulting DNA molecule is what is known as a recombinant DNA molecule.

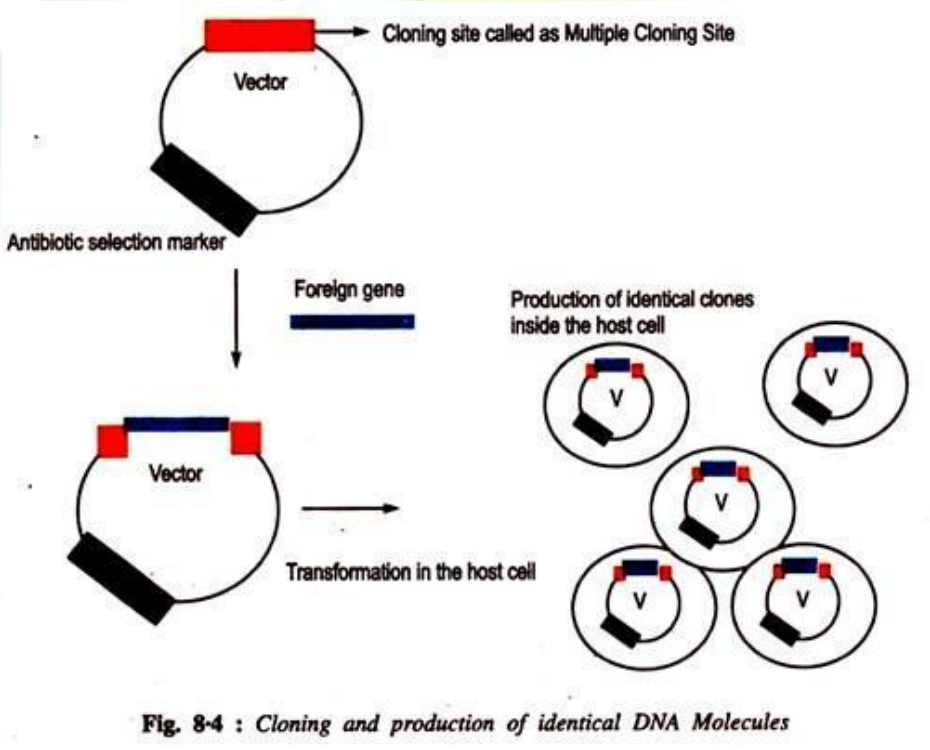
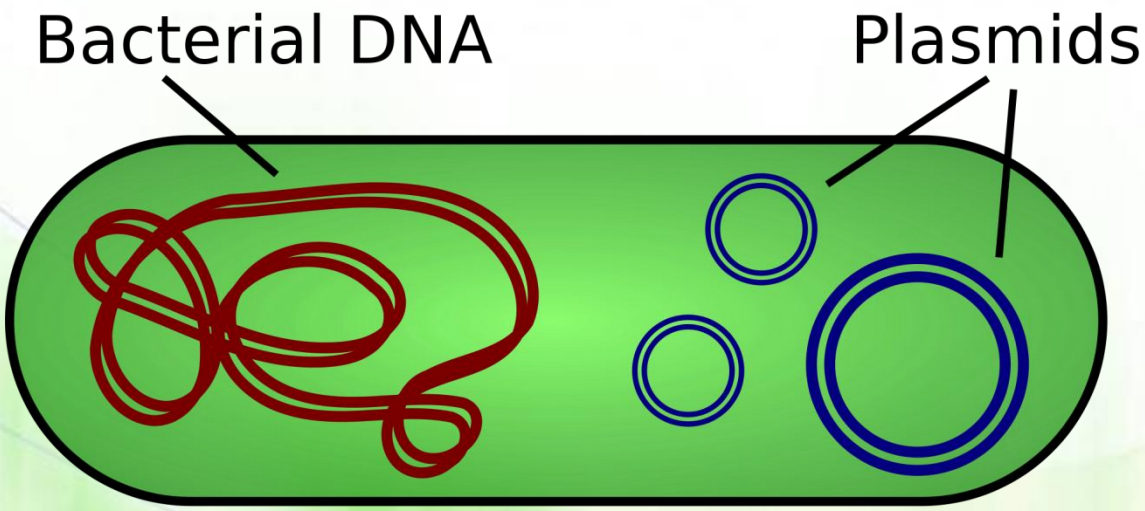


Fig. 8-4 : *Cloning and production of identical DNA Molecules*

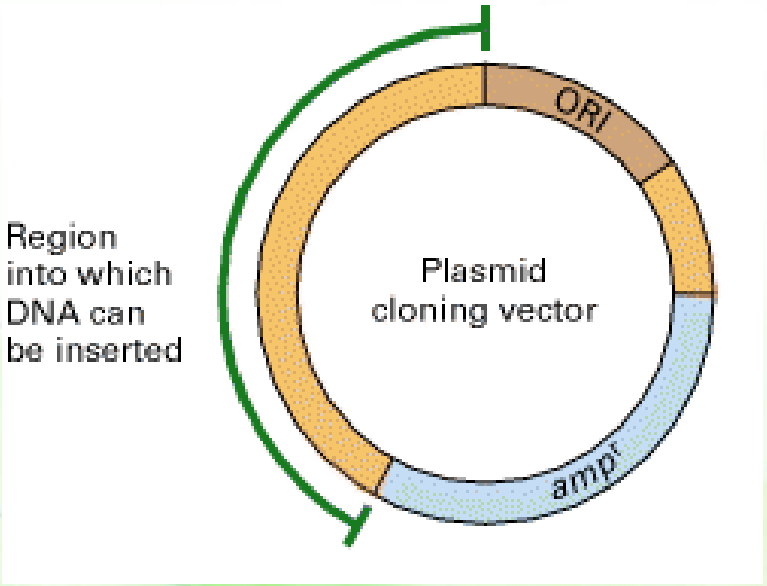
Using plasmids as vectors

- Bacterial plasmids are considered excellent vectors.
- These are bacterial circular DNA that is not part of the main circular DNA chromosome of the bacterium.
- A plasmid exists as a closed circle and replicates independently of the main bacterial genome.



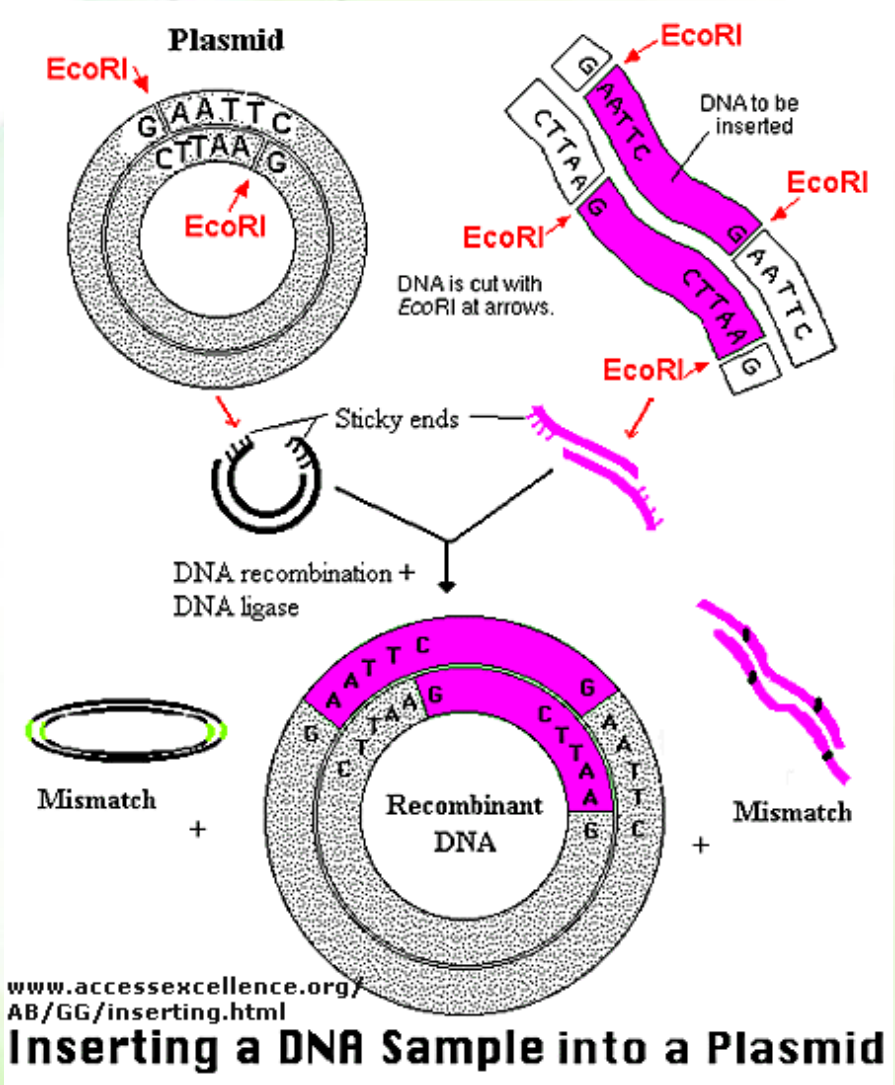
Features of plasmids

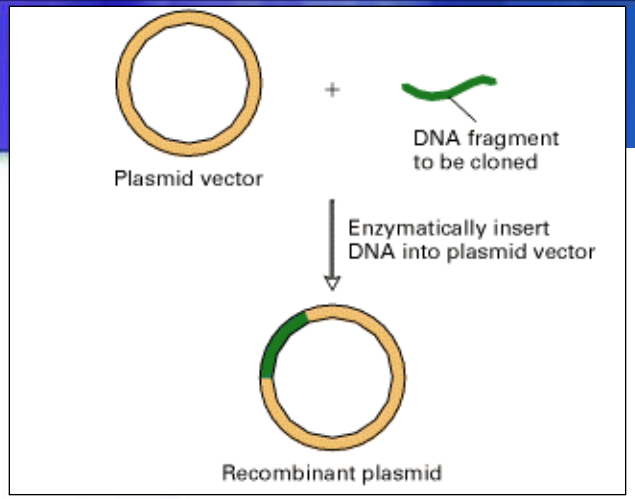
- Most plasmid vectors contain at least three essential parts required for DNA cloning:
  - Can replicate
  - Can be selected for/against by an internal drug-resistance gene (selectable marker)
  - Can inset a foreign DNA fragment

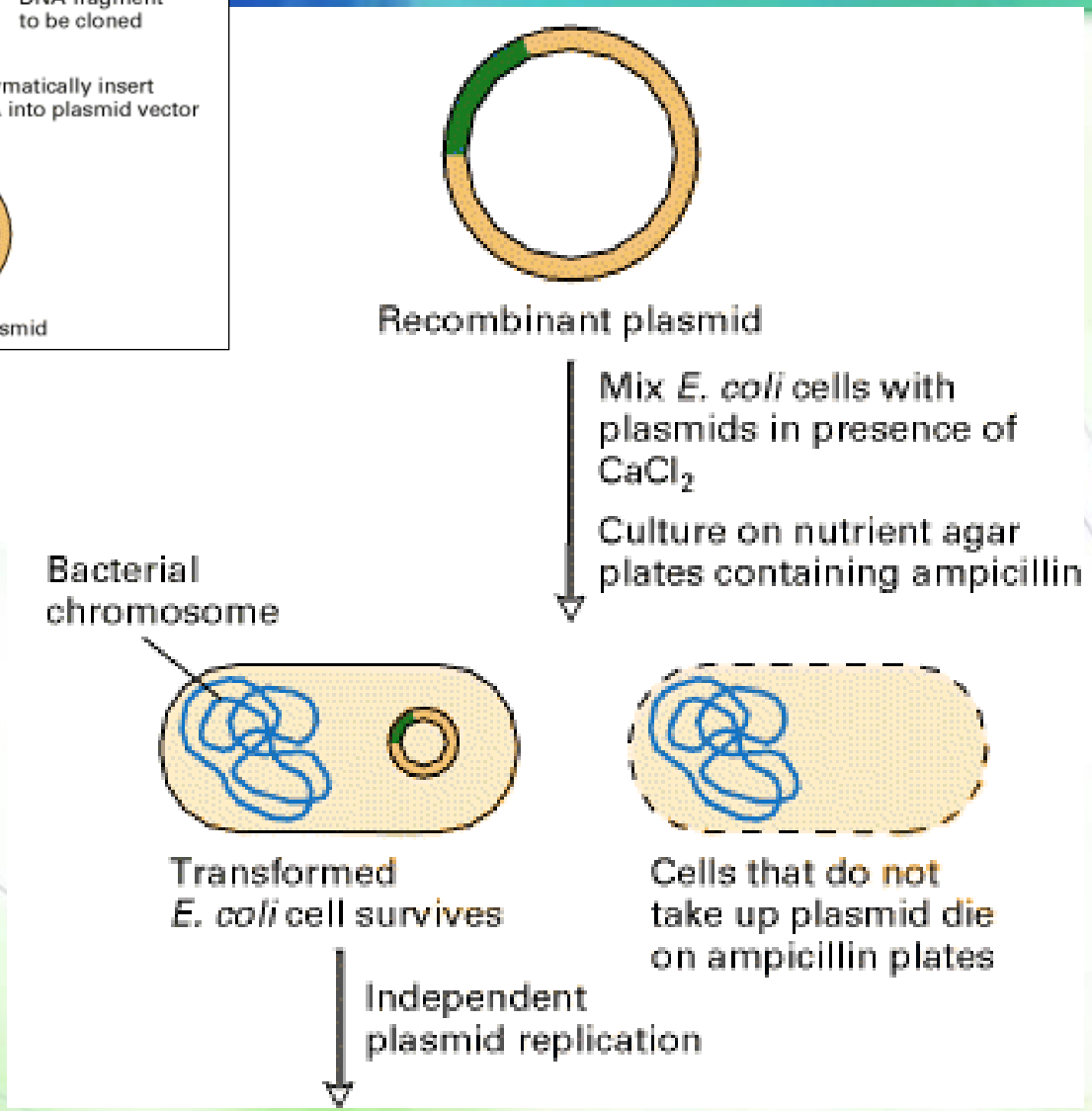


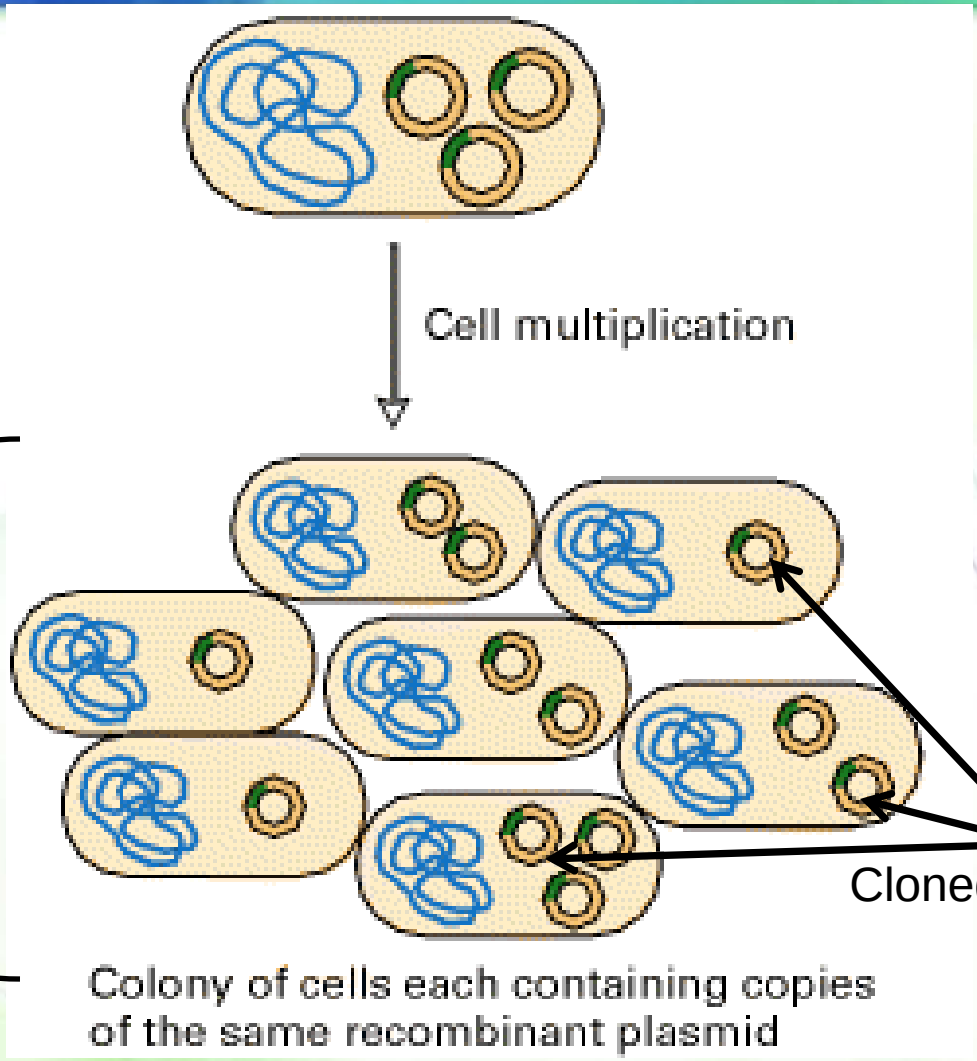
Making of recombinant DNA

- Both DNA fragments (the DNA to be cloned and a vector) are cut by the same restriction endonuclease that makes sticky-ended DNA fragments
- When mixed, they will bind to each other




Plasmid vector
+
DNA fragment
to be cloned
Enzymatically insert
DNA into plasmid vector
Recombinant plasmid


Recombinant plasmid
Mix E. coli cells with plasmids in presence of $CaCl_2$
Culture on nutrient agar plates containing ampicillin
Bacterial chromosome
Transformed E. coli cell survives
Cells that do not take up plasmid die on ampicillin plates
Independent plasmid replication

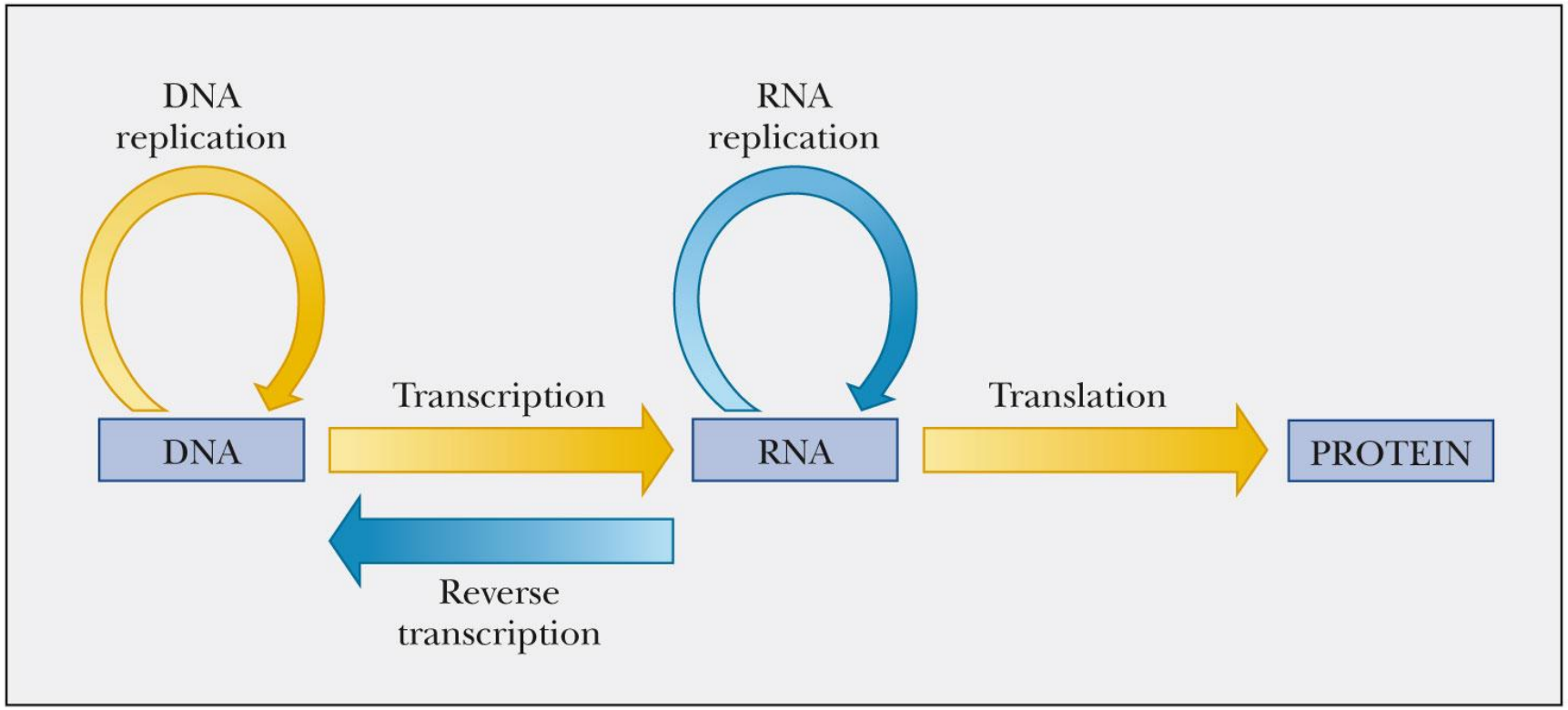

Cell multiplication
Cell clones
Cloned DNA
Colony of cells each containing copies of the same recombinant plasmid

DNA replication
a general mechanism

Resources

- This lecture
- Cooper, pp. 191-207

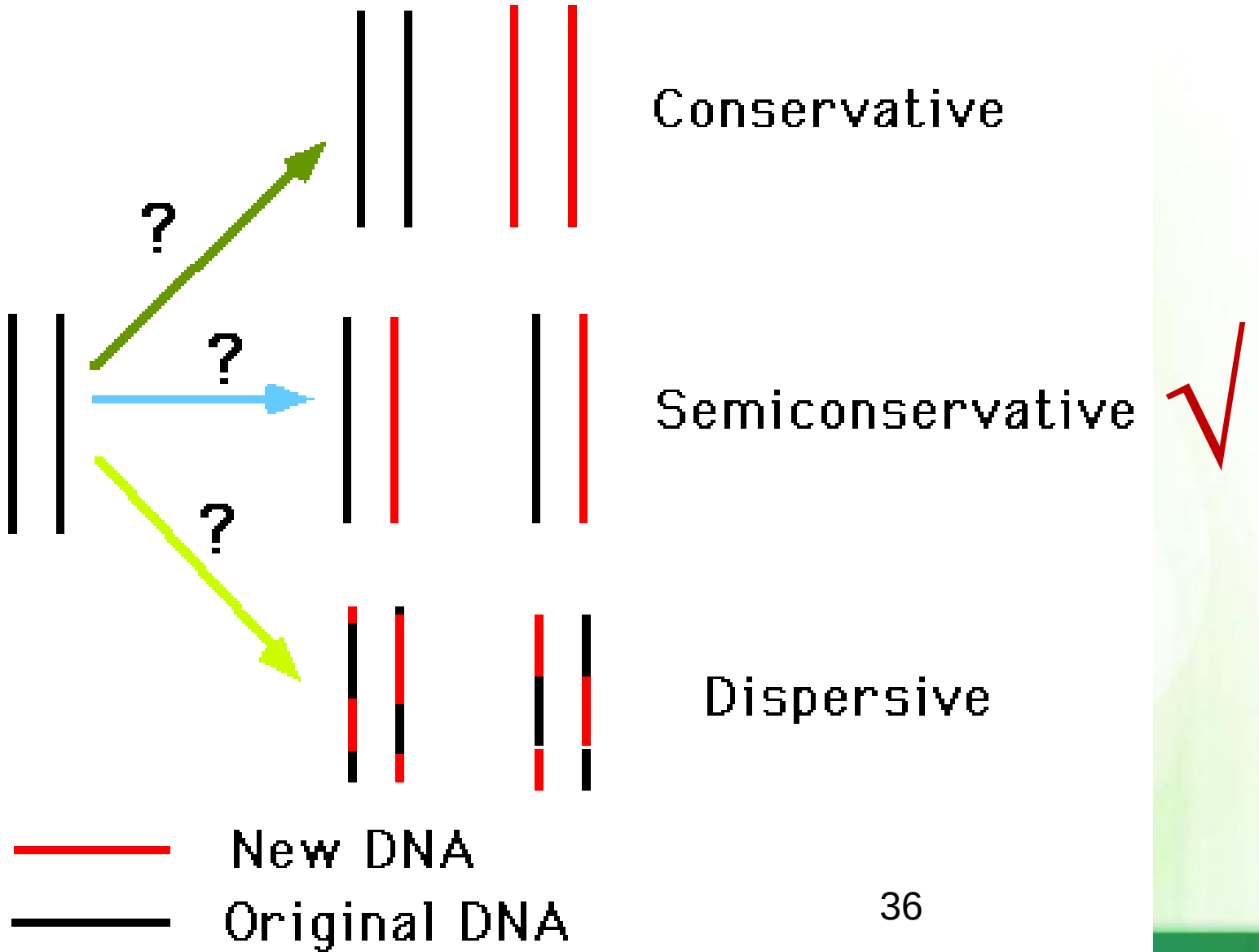
Transfer of molecular information



Some basic information

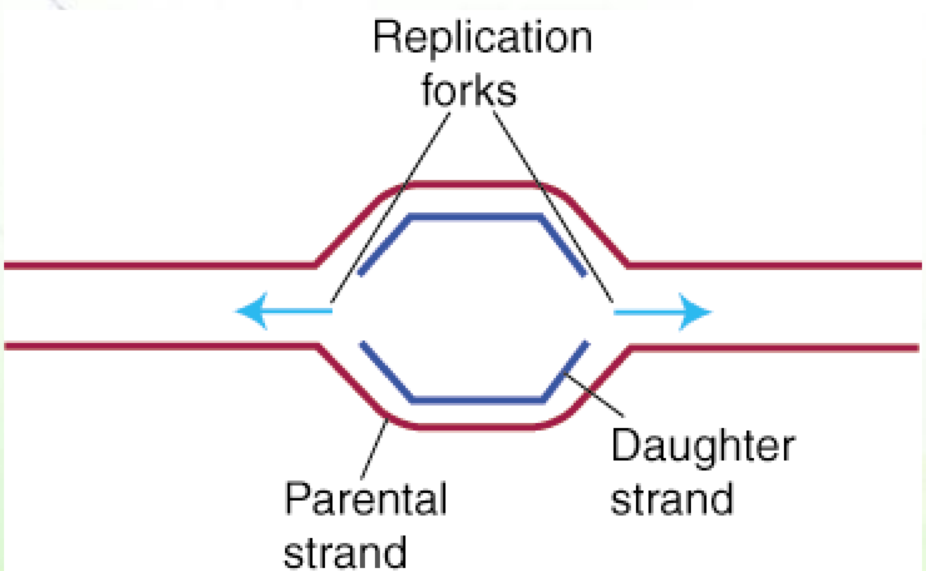
- The entire DNA content of the cell is known as genome
- DNA is organized into chromosomes
- Bacterial genome: usually one and circular chromosome
- Eukaryotic genome: multiple, linear chromosomes complexed with proteins known as histones

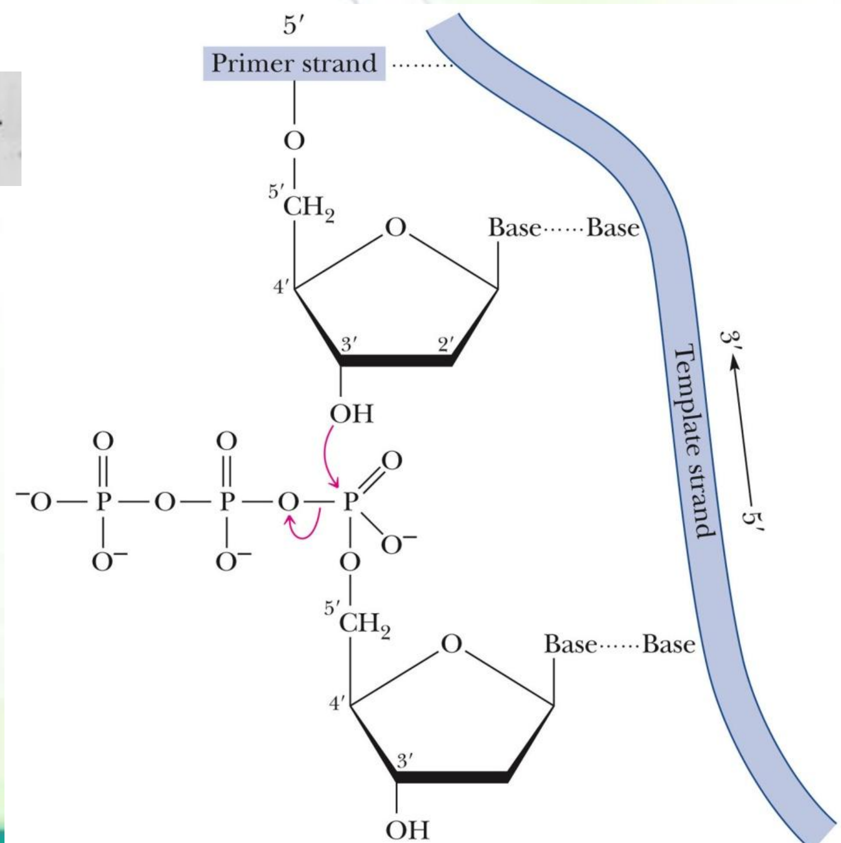
Different suggestions on possible mode of DNA replication



Bidirectionally…speaking

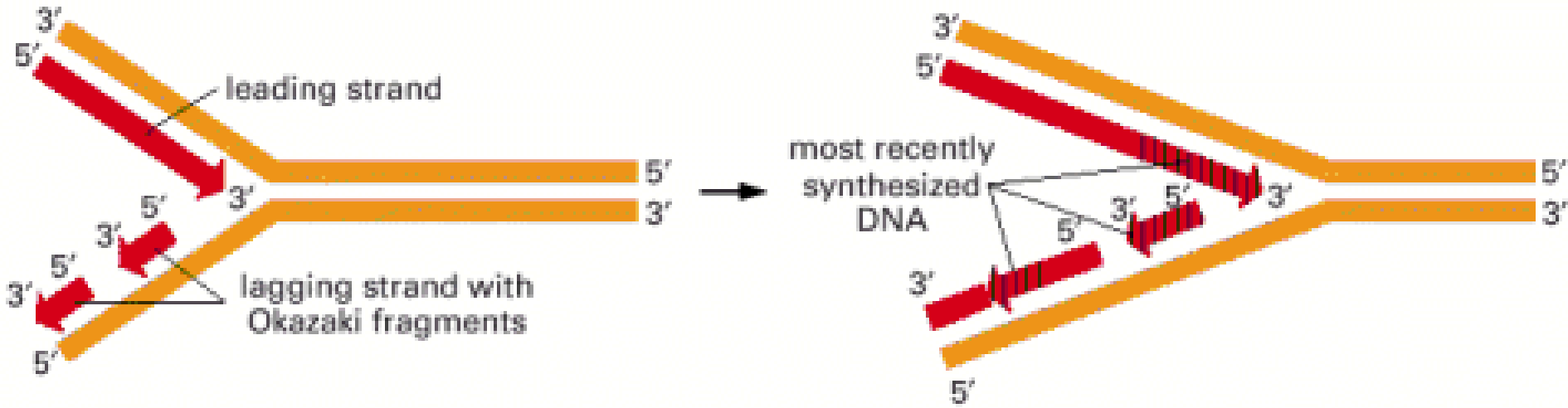
- Replication moves progressively along the parental DNA double helix bidirectionally.
- Because of its Y-shaped structure, this active region is called a replication fork.





New DNA (long vs short)

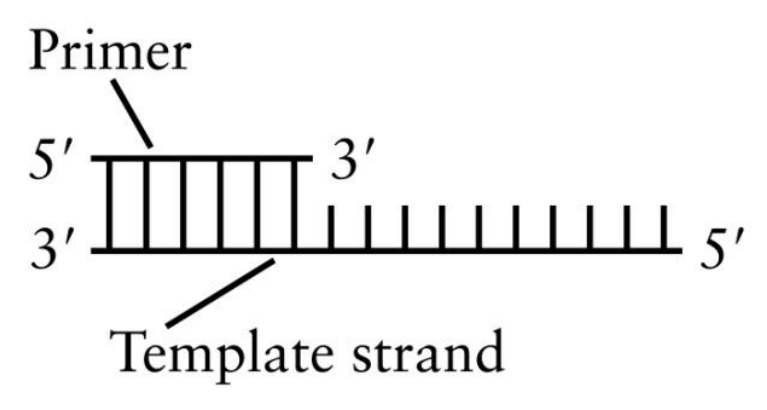
- A long strand and shorter pieces (Okazaki fragments) of DNA are present at the growing replication fork

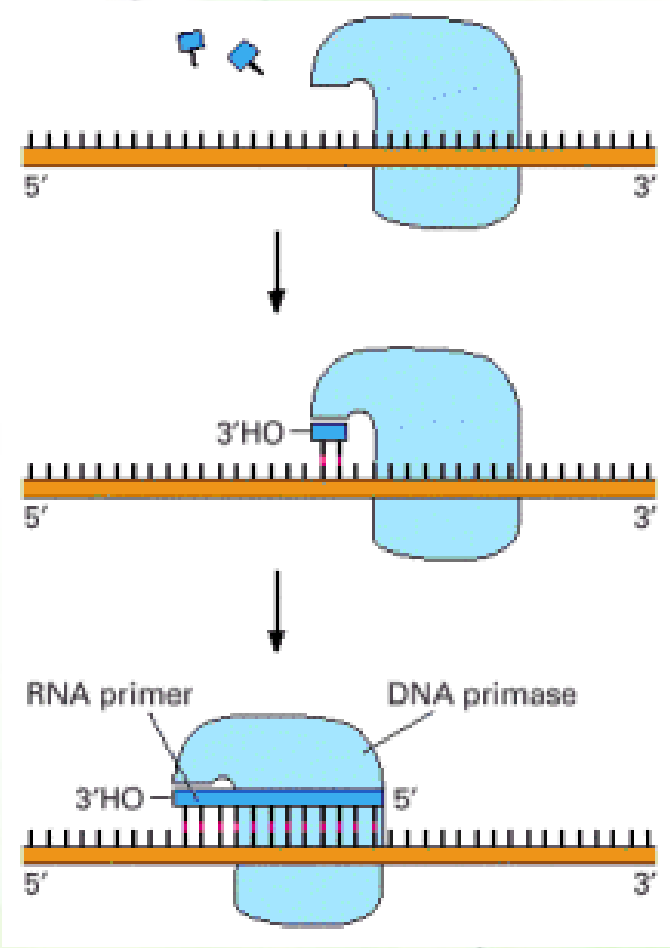


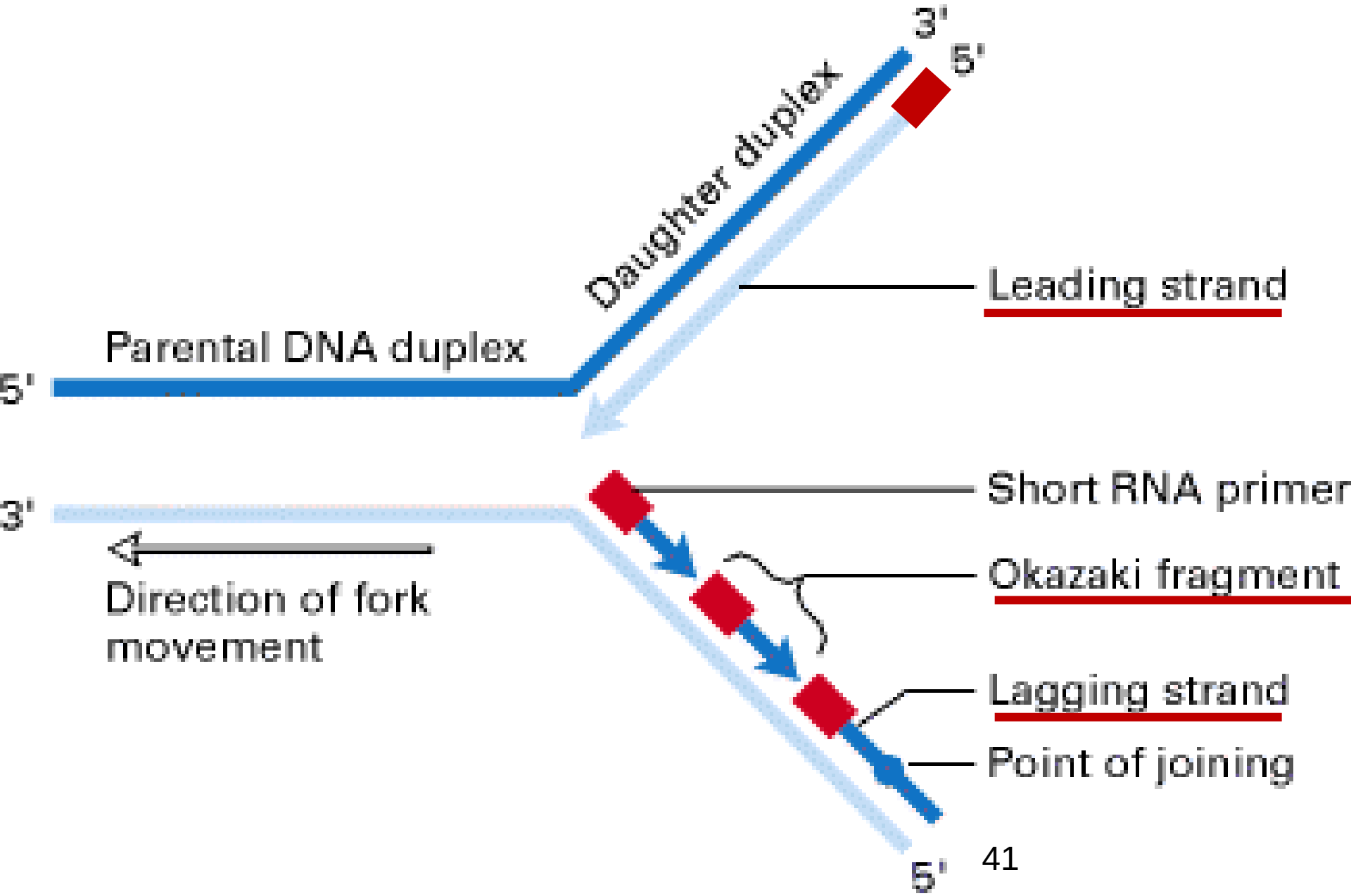
Components of DNA replication

RNA primer

- In order for the DNA polymerase to initiate replication, it requires a RNA primer, to be added first complementary to the DNA template
- It is synthesized by a primase

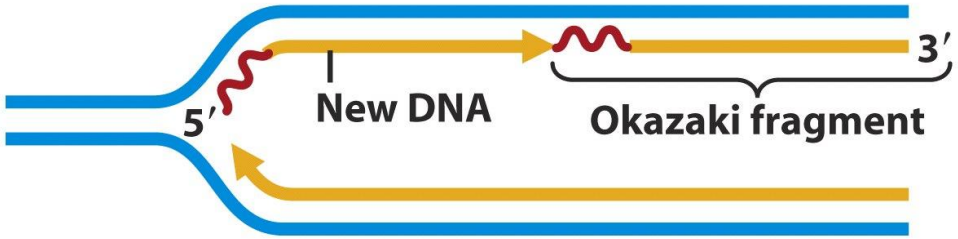




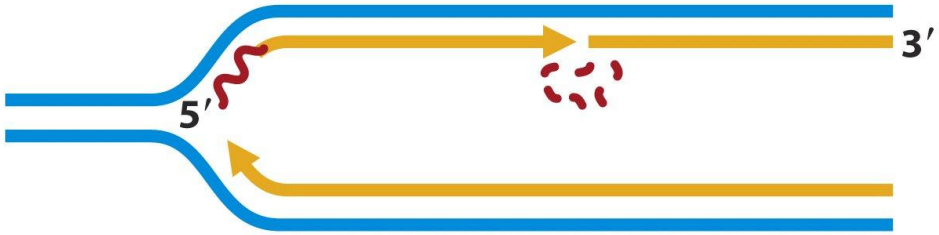

3'
5'
Daughter duplex
Leading strand
Parental DNA duplex
5'
3'
Short RNA primer
Okazaki fragment
Direction of fork movement
Lagging strand
Point of joining
5'

1. Primase synthesizes short RNA oligonucleotides (primer) copied from DNA.

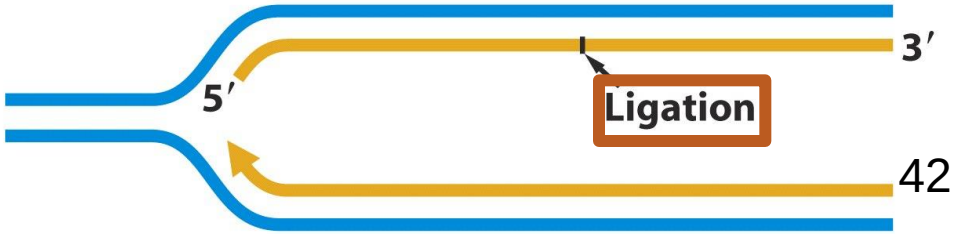
2. DNA polymerase III elongates RNA primers with new DNA.



3. DNA polymerase I removes RNA at 5′ end of neighboring fragment and fills gap.



4. DNA ligase connects adjacent fragments.



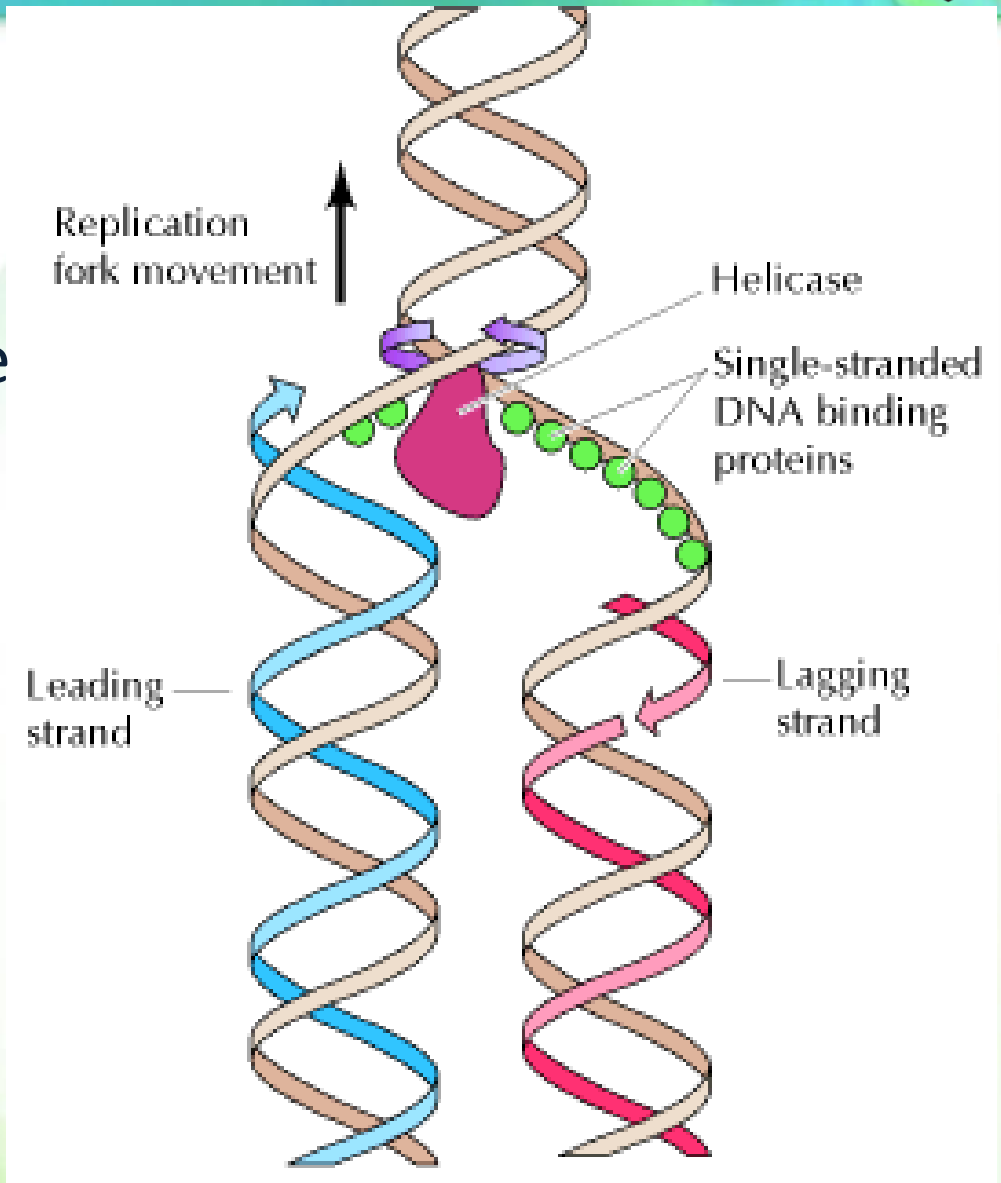
DNA helicases and SSB proteins

- For DNA synthesis to proceed, the DNA double helix must be opened up ahead of the replication fork

- Opening up the DNA is done by two types of protein contribute to this process
  - DNA helicases
  - single-strand DNA-binding proteins

DNA helicases

- DNA helicases use ATP to open up the double helical DNA as they move along the strands.
- In bacteria, helicases form a complex with the primase called primosome.



Single-strand DNA-binding (SSB) proteins

- Single-strand DNA-binding (SSB) proteins bind tightly to exposed single-stranded DNA strands without covering the bases, which remain available for templating.

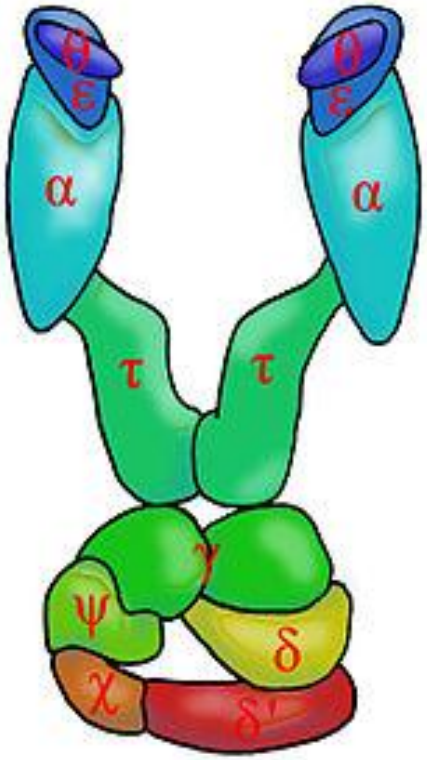
- These proteins:
  - prevent the formation of the short hairpin structures
  - protect single-stranded DNA from being degraded
  - aid helicases by stabilizing the unwound, single-stranded conformation

DNA polymerases in prokaryotes

- DNA polymerase III: DNA polymerization at the growing fork in E. coli
  - The complex of primosome and polymerase is known as replisome
- DNA polymerase I:
  - 5‘-to-3' exonuclease activity (removal of RNA primer) of each Okazaki fragment.
  - Fills in the gaps between the lagging-strand fragments.
  - DNA repair
- DNA polymerase II, IV, and V : DNA repair

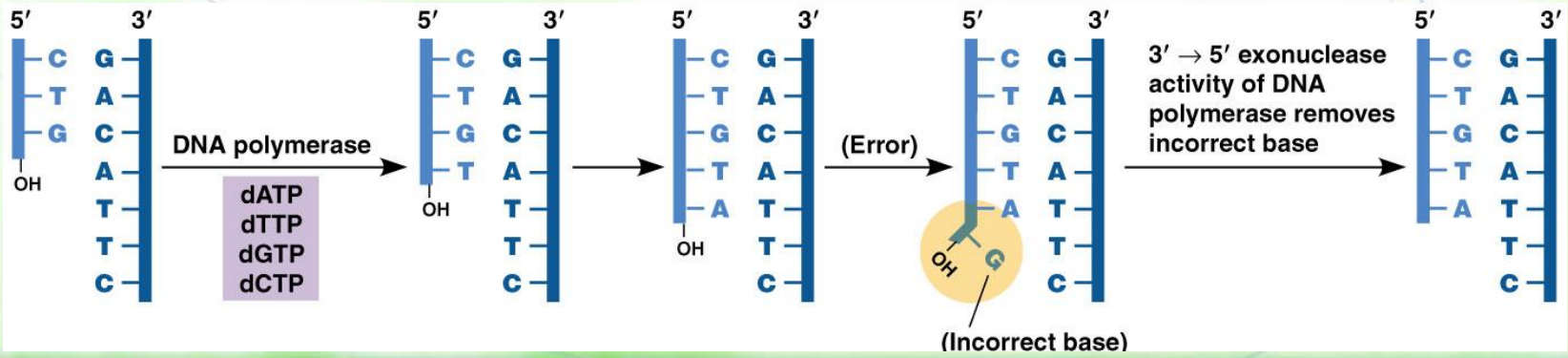
DNA polymerase III

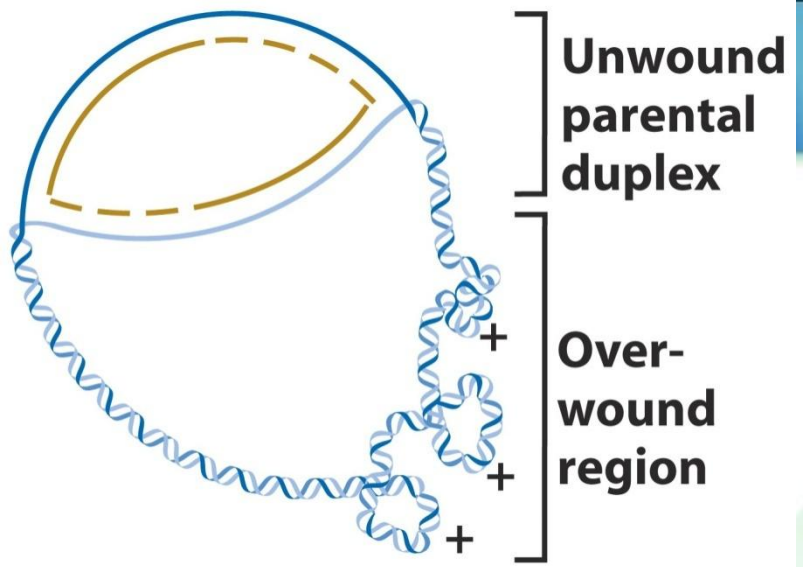
- The DNA polymerase III is a very large protein composed of 10 different polypeptides.
- The core polymerase is composed of three subunits:
  - α subunit contains the active site for nucleotide addition.
  - ε subunit is a 3’-to-5’ exonuclease that removes incorrectly added (mispaired) nucleotides from the end of the growing chain.

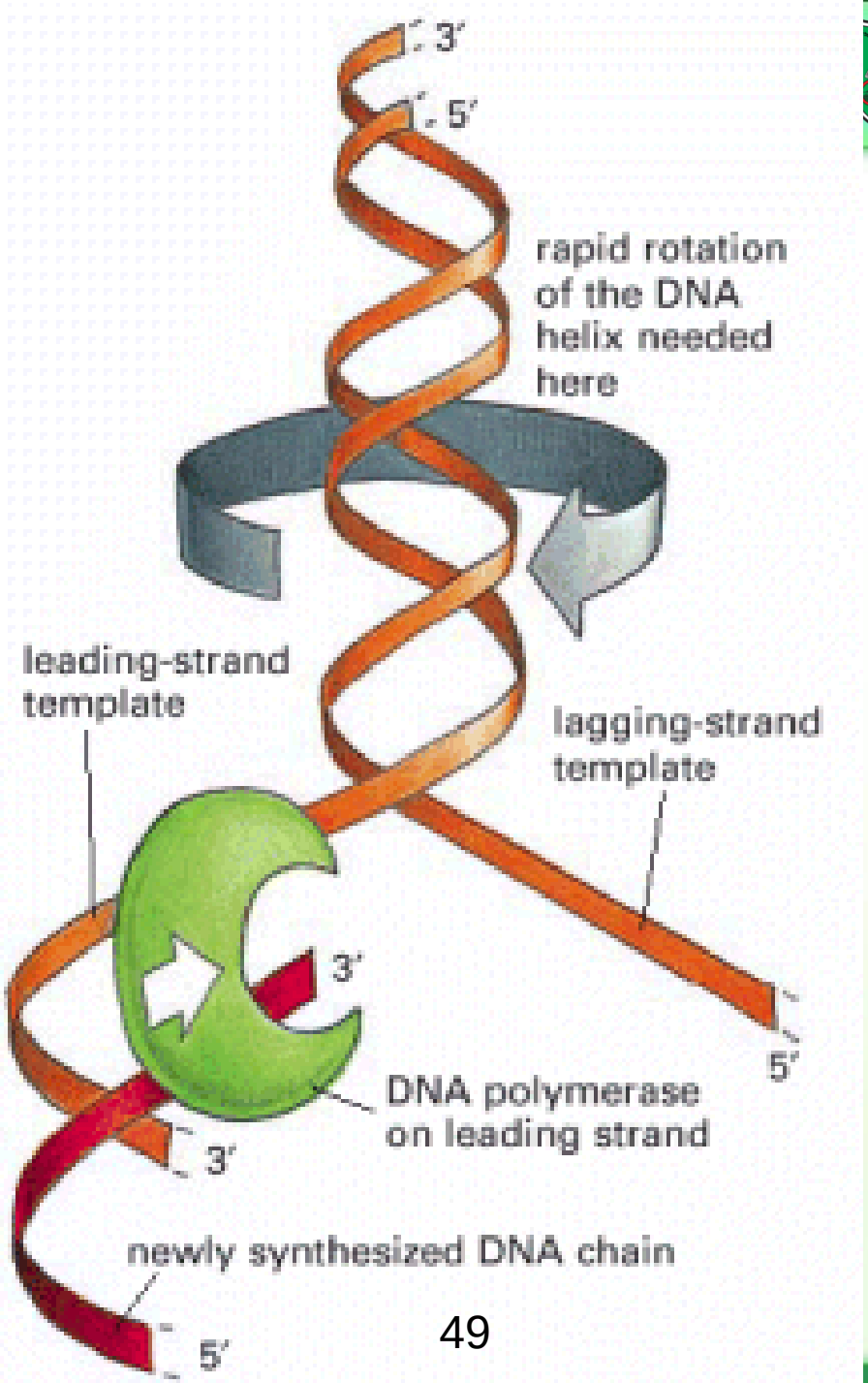


How accurate is DNA replication?

- The frequency of errors during replication is only one incorrect base per 10^9 to 10^{10} nucleotides incorporated
- Why is fidelity high?
  - Hydrogen base-pairing is highly stable between G and C and between A and T. So, the DNA polymerase can catalyze the formation of phosphodiester bonds when the right hydrogen bonding takes place between the correction bases.
  - Proofreading mechanism (a 3'→5' exonuclease activity)- Remember ε subunit of RNA pol III

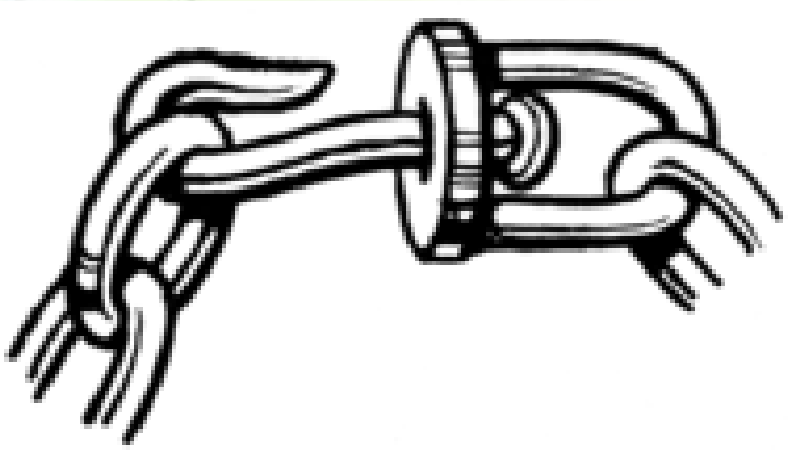


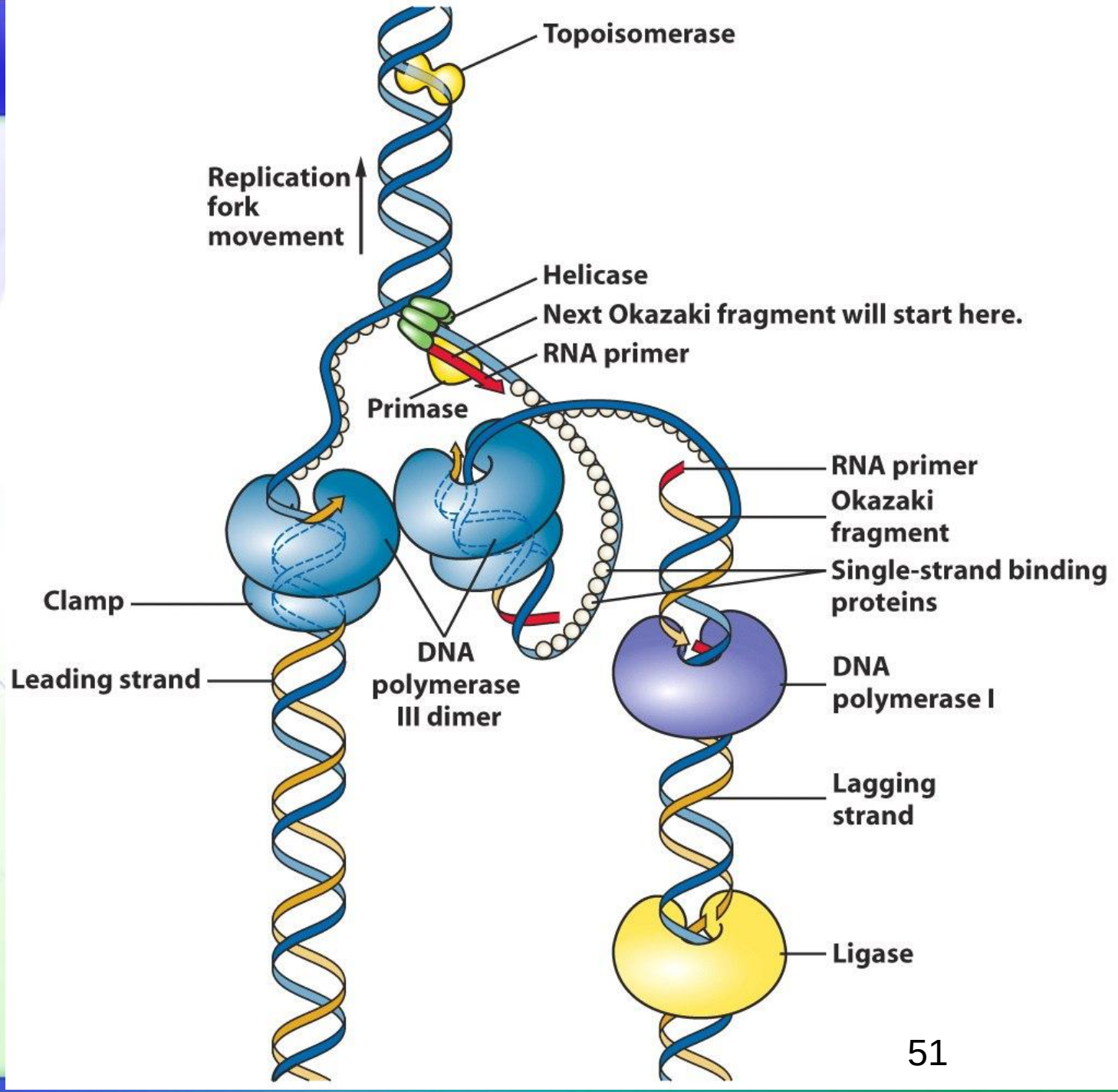

Unwound parental duplex
Over-wound region
+
+
+


3′
5′
rapid rotation of the DNA helix needed here
leading-strand template
lagging-strand template
3′
5′
DNA polymerase on leading strand
3′
newly synthesized DNA chain
5′

DNA topoisomerases

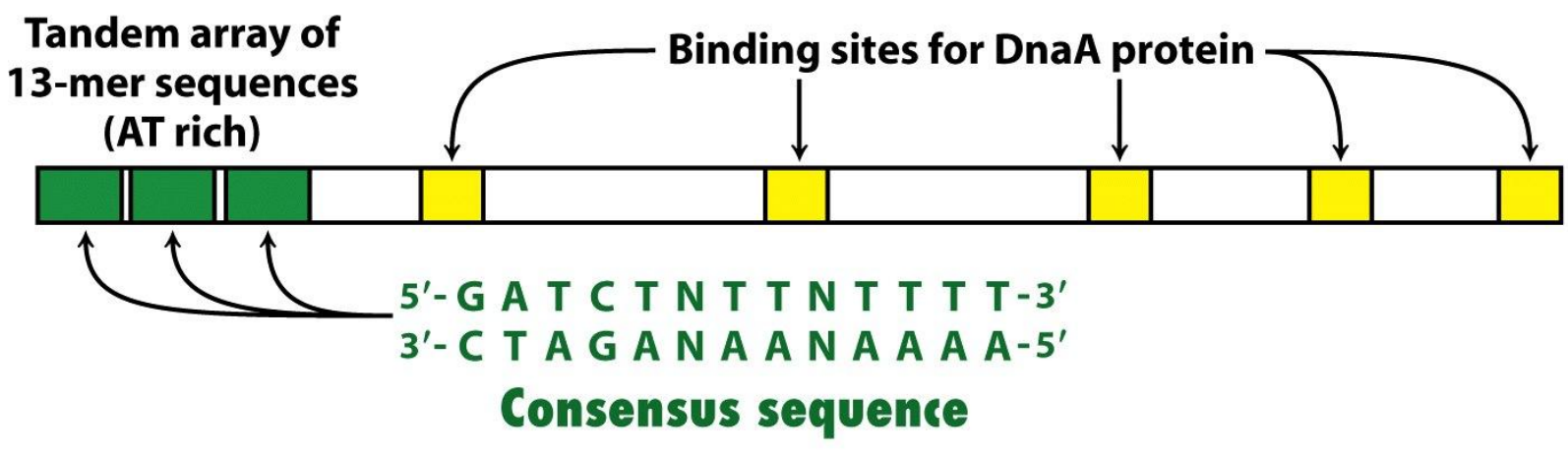
- A swivel is formed in the DNA helix by proteins known as DNA topoisomerases
- A DNA topoisomerase breaks then re-forms phosphodiester bonds in a DNA strand.
- Topoisomerase I produces an transient single-strand break (or nick)
  - ATP-independent
- Topoisomerase II is responsible for untangling chromosomes by making a transient double-strand break
  - also known as gyrase in bacteria
  - ATP-dependent




Topoisomerase
Replication fork movement
Helicase
Next Okazaki fragment will start here.
RNA primer
Primase
RNA primer
Okazaki fragment
Single-strand binding proteins
Clamp
Leading strand
DNA polymerase III dimer
DNA polymerase I
Lagging strand
Ligase

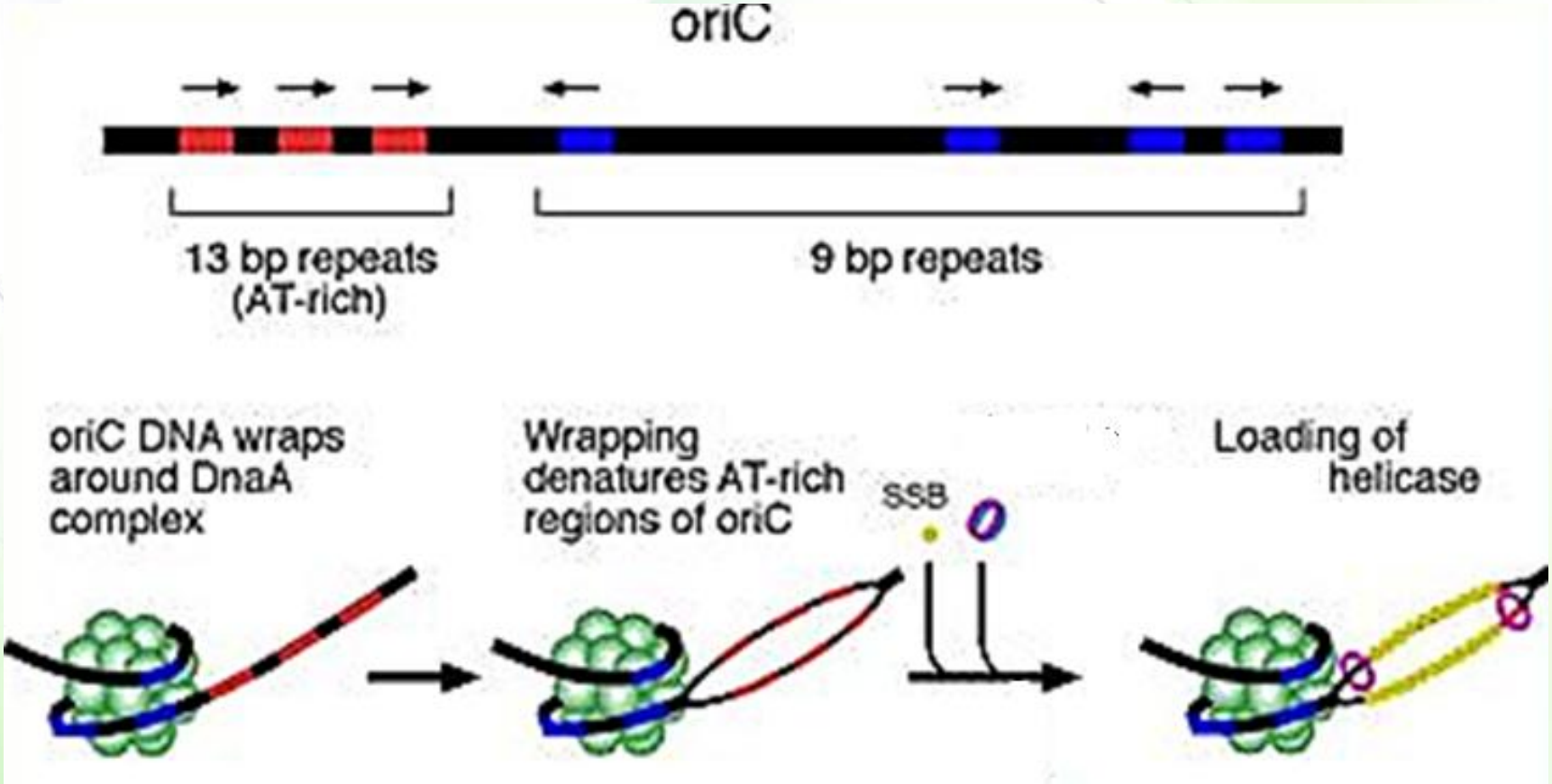
Replication Origin in bacteria

- Bacterial replication starts at a origin known as origin of replication (OriC)
- oriC regions contain repetitive 9-bp and AT-rich 13-bp sequences
- 9-mer: binding sites for the DnaA protein
- 13-mers: AT-rich region
  - facilitates separation of the double strand DNA



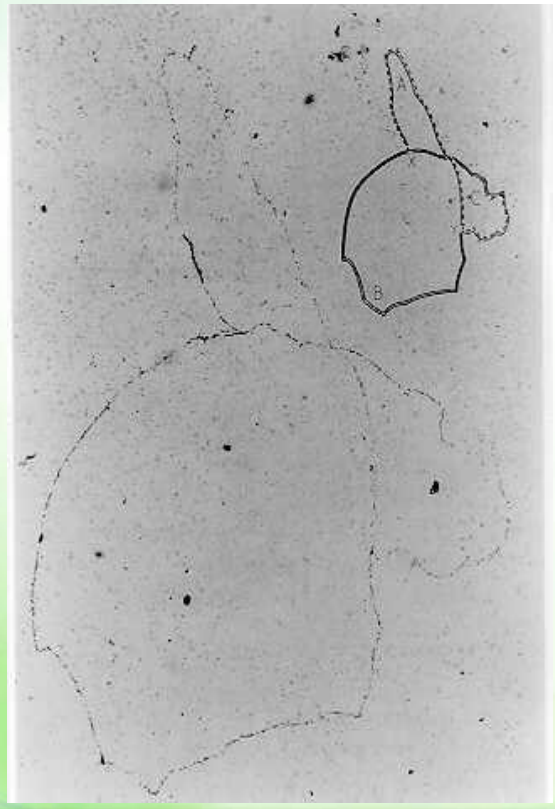
Possible mechanism

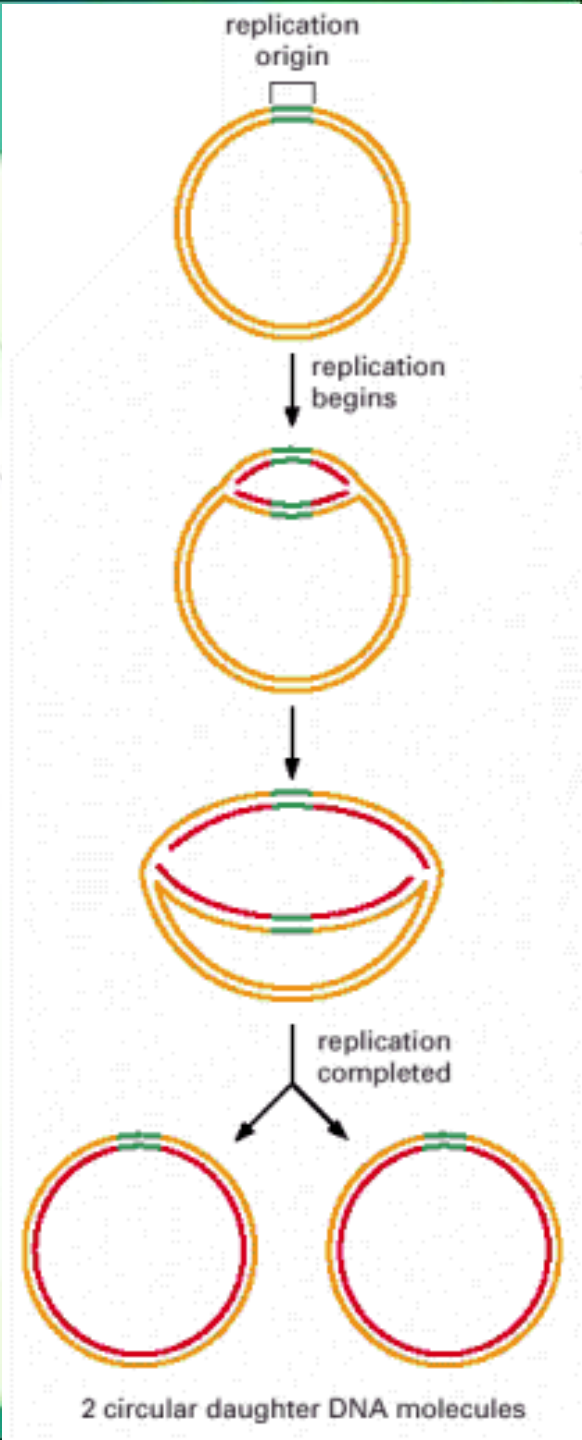
- When DnaA protein binds to 9-mers, it applies stress on the AT-rich region resulting in DNA "melting“.



Two replication forks (bacteria)

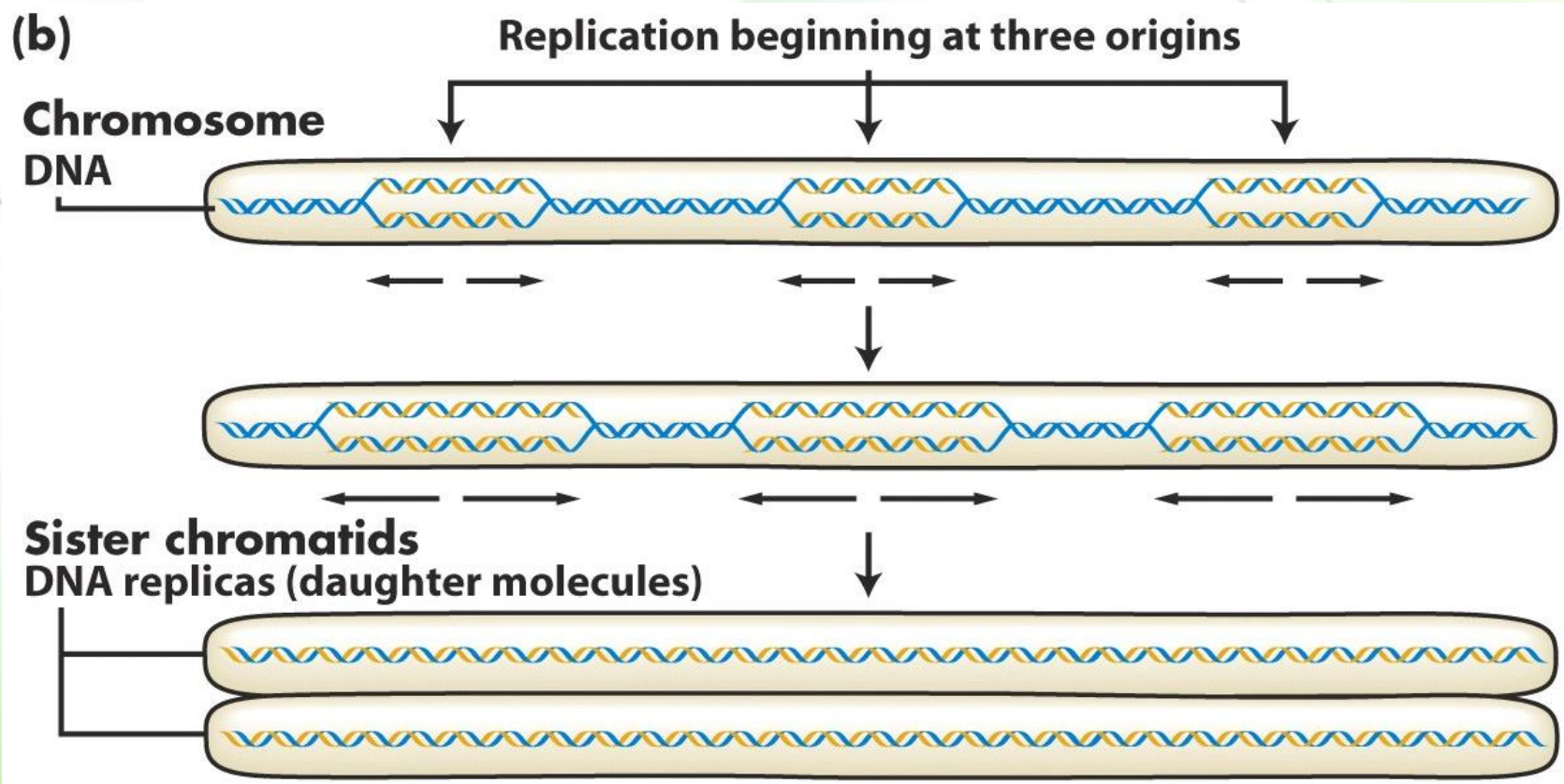
- The two replication forks proceed in opposite directions until they meet up roughly halfway around the chromosome.





Origins of replication in human genome

- An average human chromosome may have several hundred replicators (origins of replication).



DNA polymerase in eukaryotes

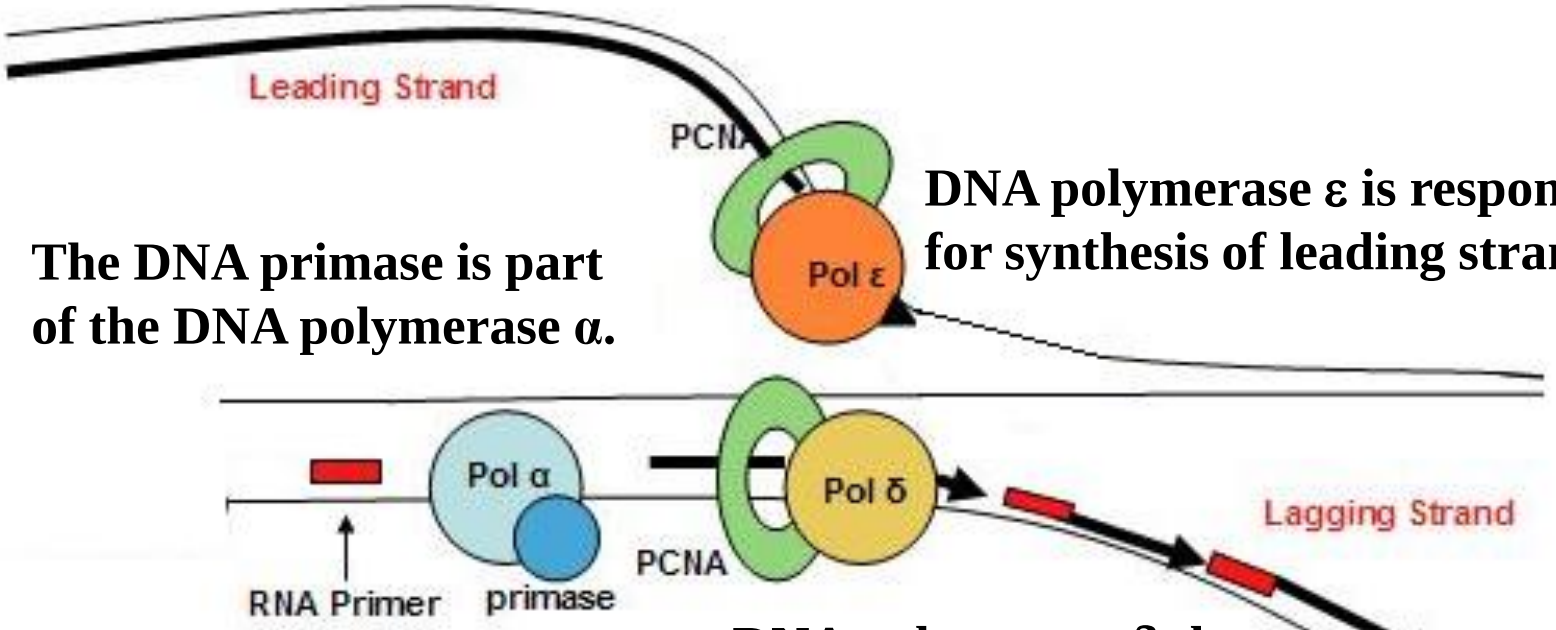
- Eukaryotic cells contain 9 DNA polymerases; most of them for DNA repair.

TABLE 10.4

The Biochemical Properties of Eukaryotic DNA Polymerases

	α	δ	ε	β	γ
Mass (kDa)					
Native	>250	170	256	36-38	160-300
Catalytic core	165-180	125	215	36-38	125
Other subunits	70, 50, 60	48	55	None	35, 47
Location	Nucleus	Nucleus	Nucleus	Nucleus	Mitochondria
Associated functions					
3' → 5' exonuclease	No	Yes	Yes	No	Yes
Primase	Yes	No	No	No	No
Properties					
Processivity	Low	High	High	Low	High
Fidelity	High	High	High	Low	High
Replication	Yes	Yes	Yes	No	Yes
Repair	No	?	Yes	Yes	No

The mechanism of replication



The DNA primase is part of the DNA polymerase α.

DNA polymerase ε is responsible for synthesis of leading strand.

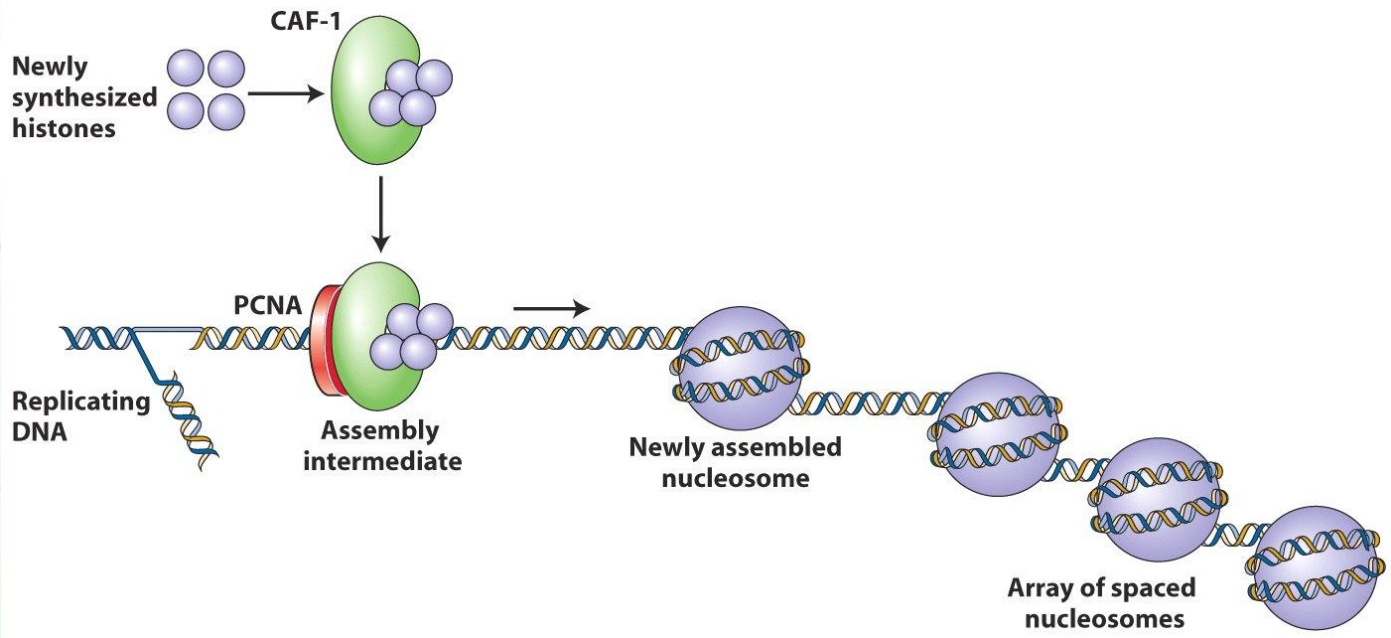
Polymerase α begins each Okazaki fragment on the lagging strand

DNA polymerase δ then synthesizes the remainder of each Okazaki fragment.

- The polymerases do not have a 5'→3' exonuclease.
  - The primer is removed by two special enzymes.
  - DNA polymerase δ then fills in the gap.

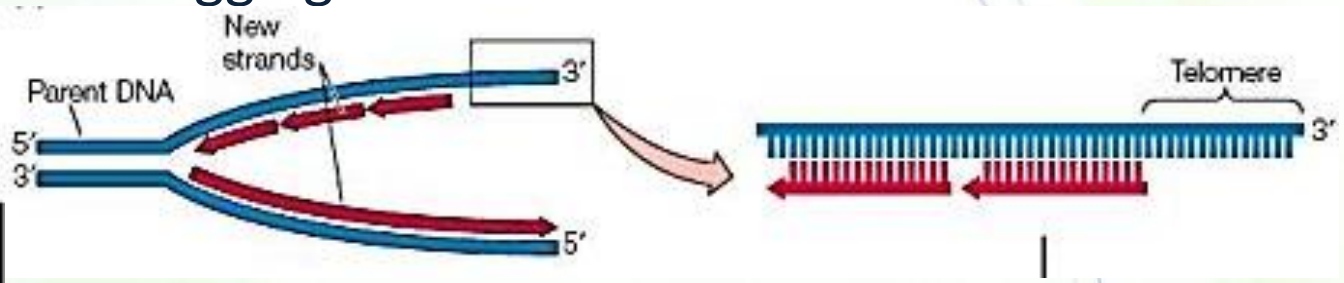
Role of chromatin

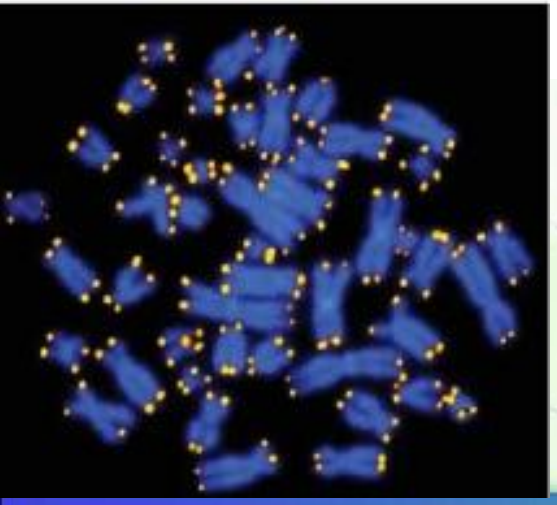
- Replication is linked to DNA packing by histones.
- DNA is freed from histones by chromatin-remodeling proteins in order for enzymes to move along the DNA.
- New histones are assembled onto the DNA behind each replication fork by chromatin assembly factors (CAFs).

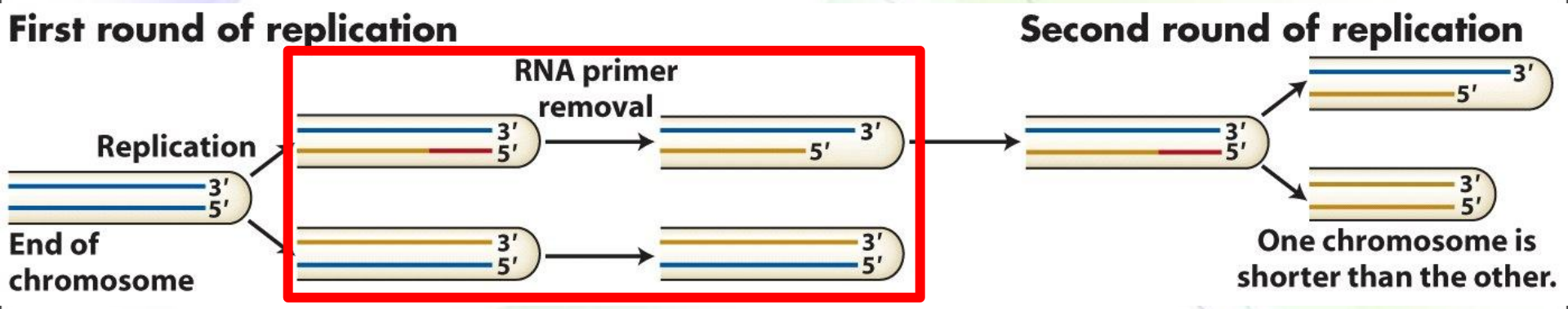


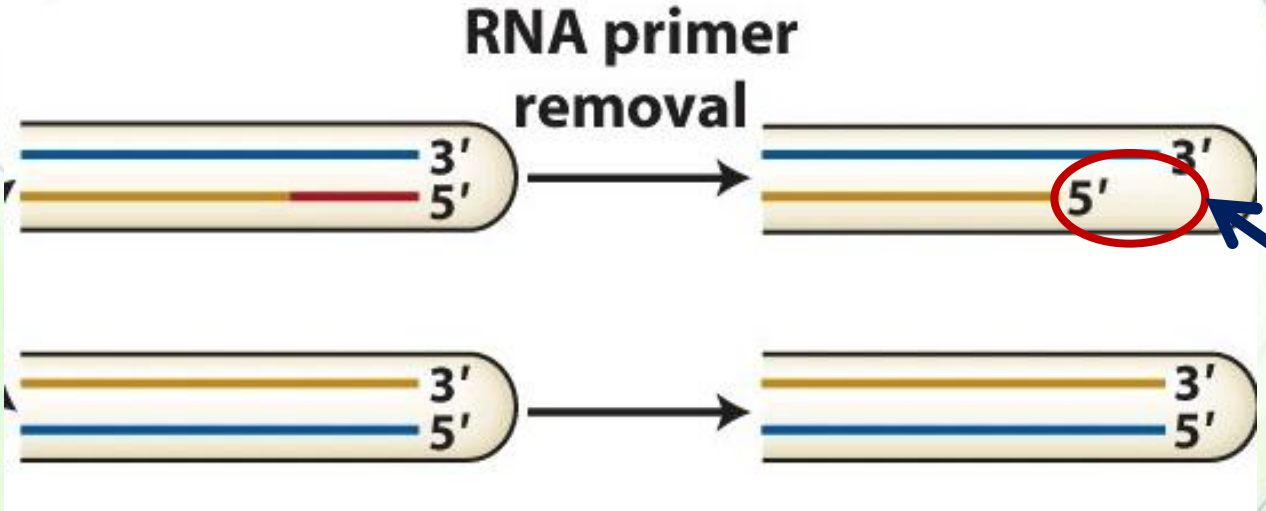
A problem in the lagging strand

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



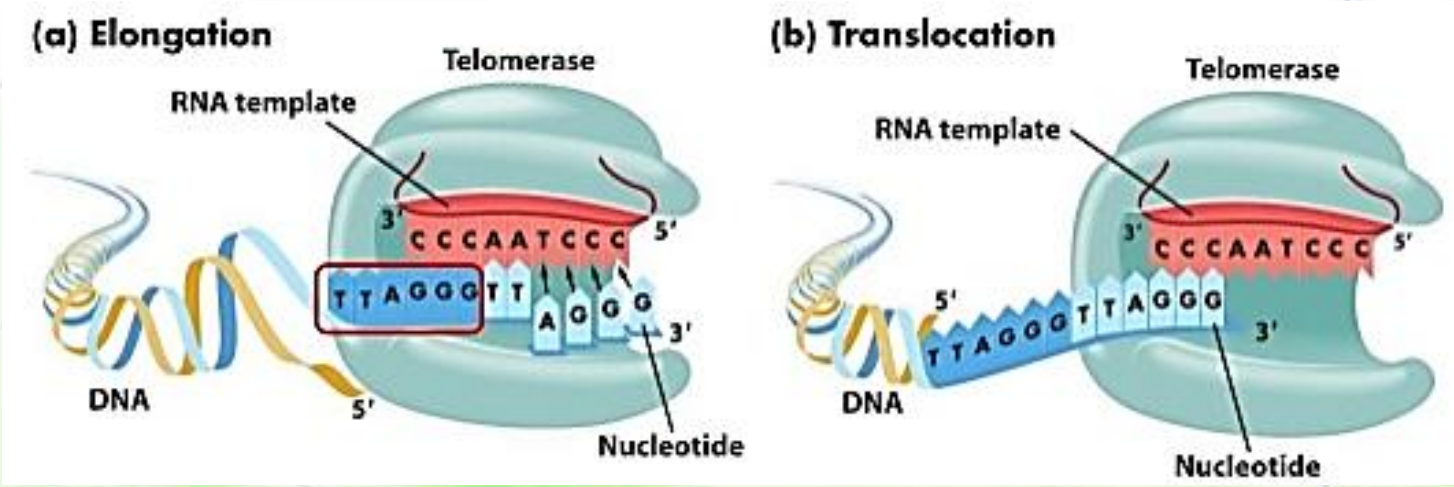


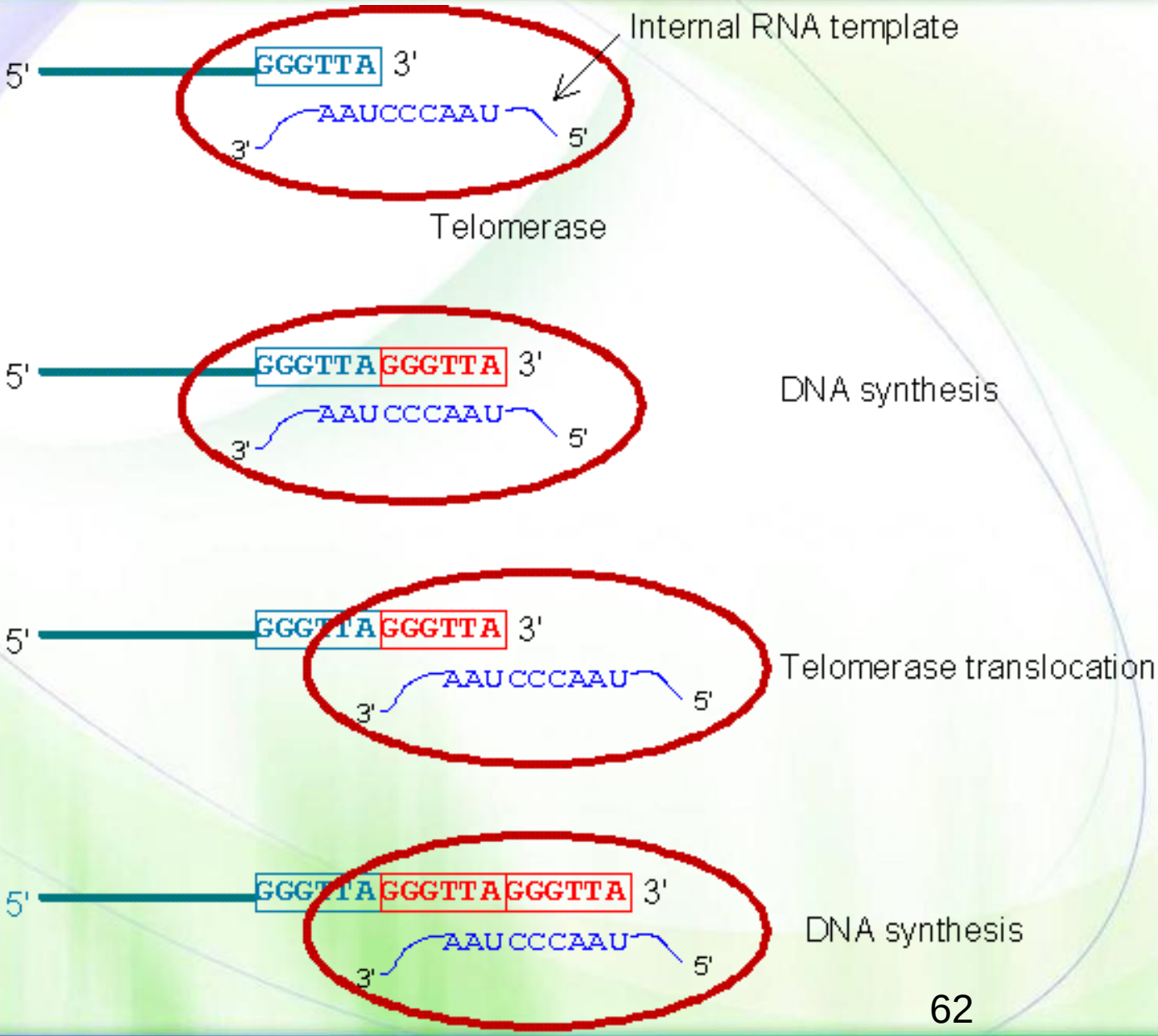

First round of replication
Second round of replication
RNA primer removal
Replication
3′
5′
End of chromosome
One chromosome is shorter than the other.


RNA primer removal
3′
5′
Primer gap

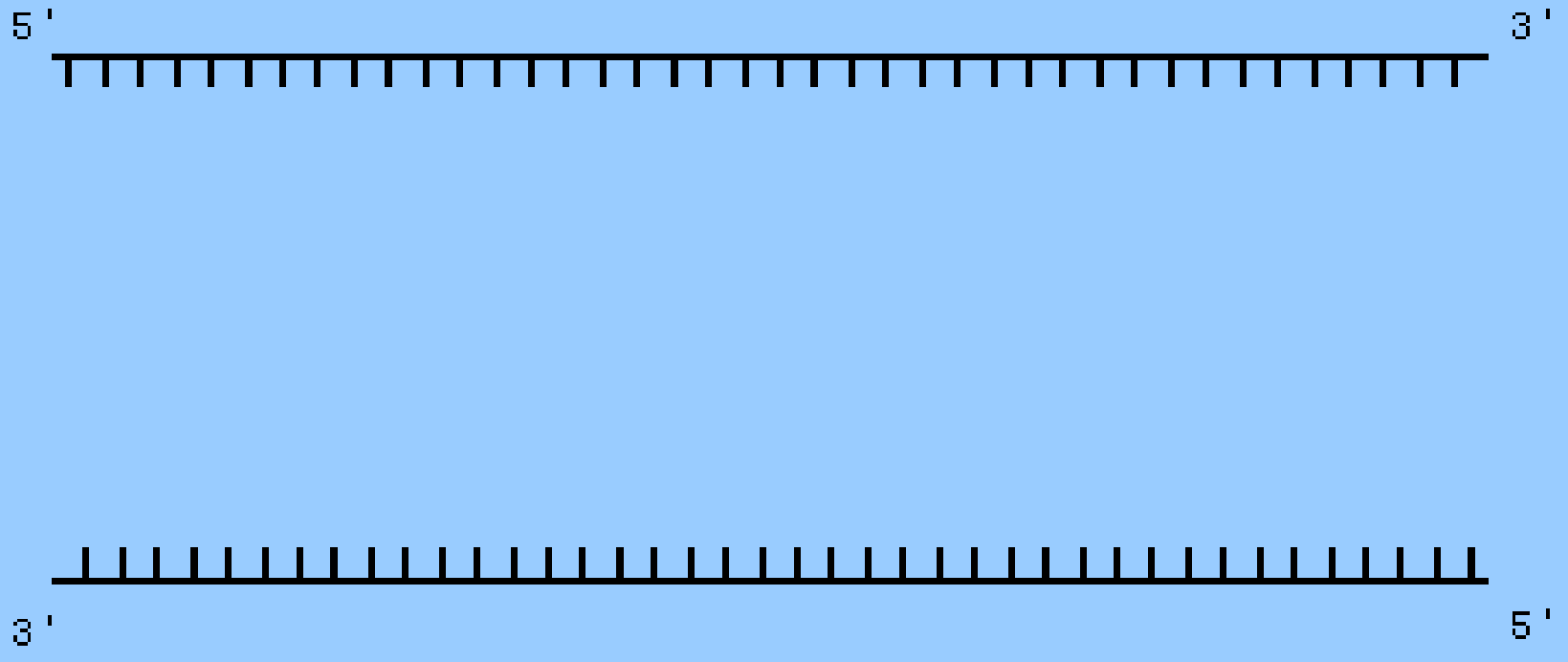
Telomerase comes to the rescue

- Telomerase prevents the progressive shortening of the lagging strand. *How*?
- Telomere DNA sequences consist of many GGGTTA repeats extending about 10,000 nucleotides.
- Telomerase recognizes the repeat sequence and elongates it in the 5'-to-3' direction using a RNA template/primer that is a component of the enzyme itself.




Internal RNA template
5'
GGGTTA 3'
AAUCCCAAU
3'
5'
Telomerase
5'
GGGTTA GGGTTA 3'
AAUCCCAAU
3'
5'
DNA synthesis
5'
GGGTTA GGGTTA 3'
AAUCCCAAU
3'
5'
Telomerase translocation
5'
GGGTTA GGGTTA GGGTTA 3'
AAUCCCAAU
3'
5'
DNA synthesis

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



How do we age?

- As we grow older, the activity of telomerase is reduced.
- An inverse relationship between age and telomeric length has been observed.
- The gradual shortening of the chromosome ends leads to cell death, and it has even been suggested that life span is determined by the length of telomeres.

Elixir of youth