

BIOINFORMATICS

(FOR COMPUTER SCIENTISTS)

MPCS56420
SPRING 2020
SESSION 2



BIOLOGY FOR DNA SEQUENCING

BIOLOGY FOR DNA SEQUENCING

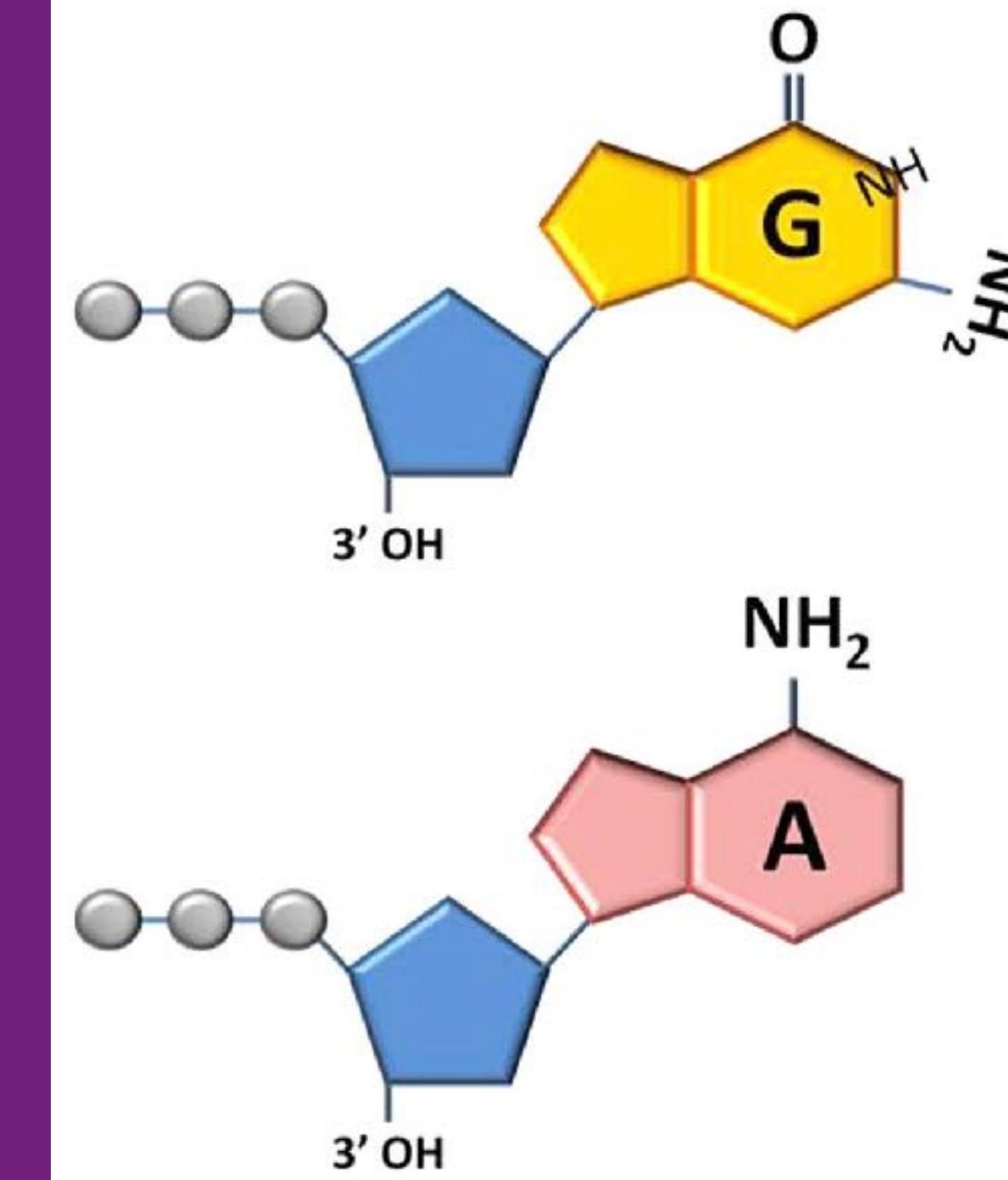
- Step 1: DNA Extraction
 - Isolate the DNA we are interested in (e.g. gene)
- Step 2: DNA Cloning
 - Amplify the DNA so we have enough to work with
- Step 3: Sequencing
 - Put it in an expensive machine and push a button



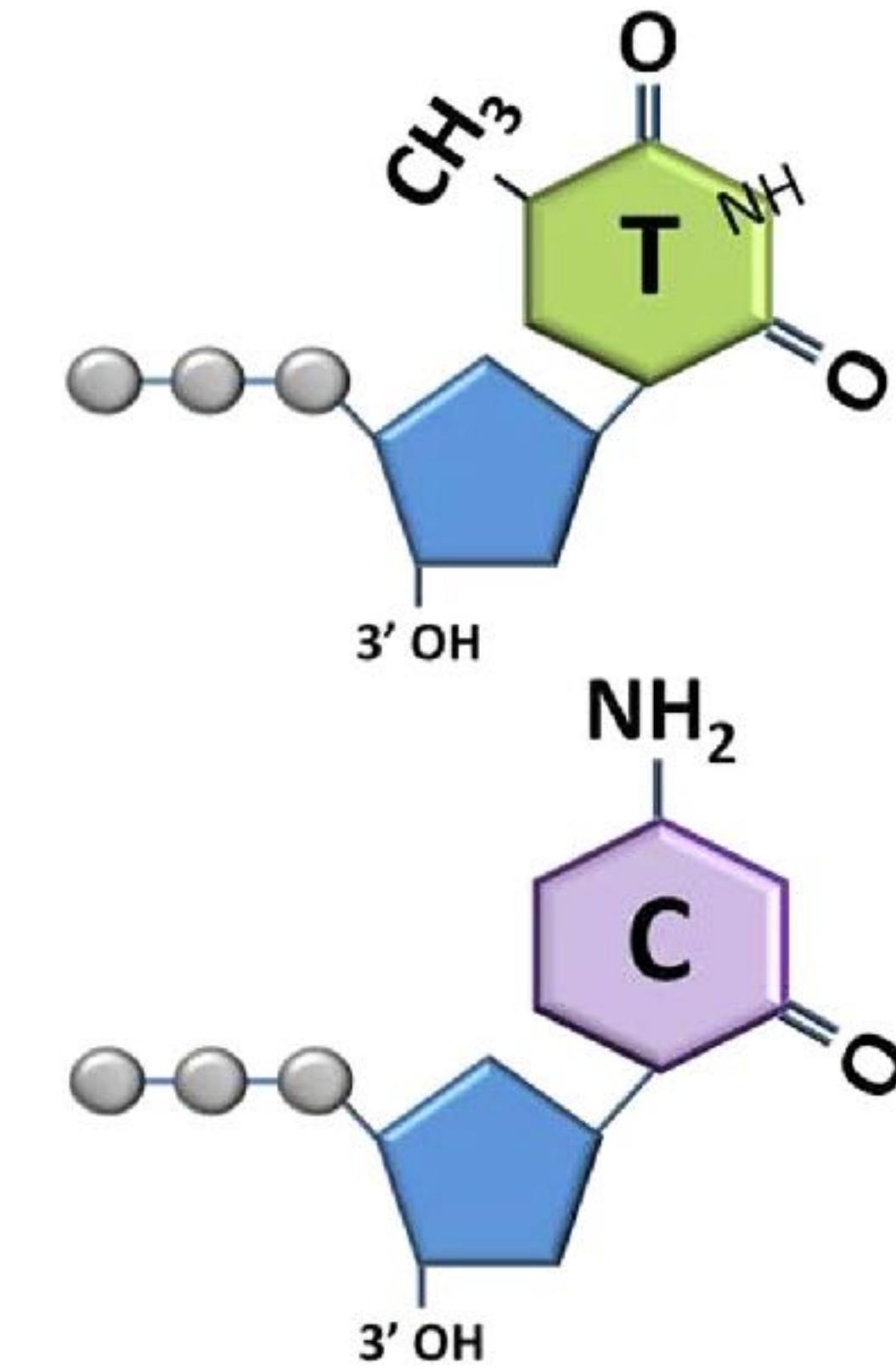
DNA SEQUENCING

IMPORTANT MOLECULES FOR SEQUENCING

- Nucleotides
 - Components in DNA



Purine
Deoxyribonucleotides

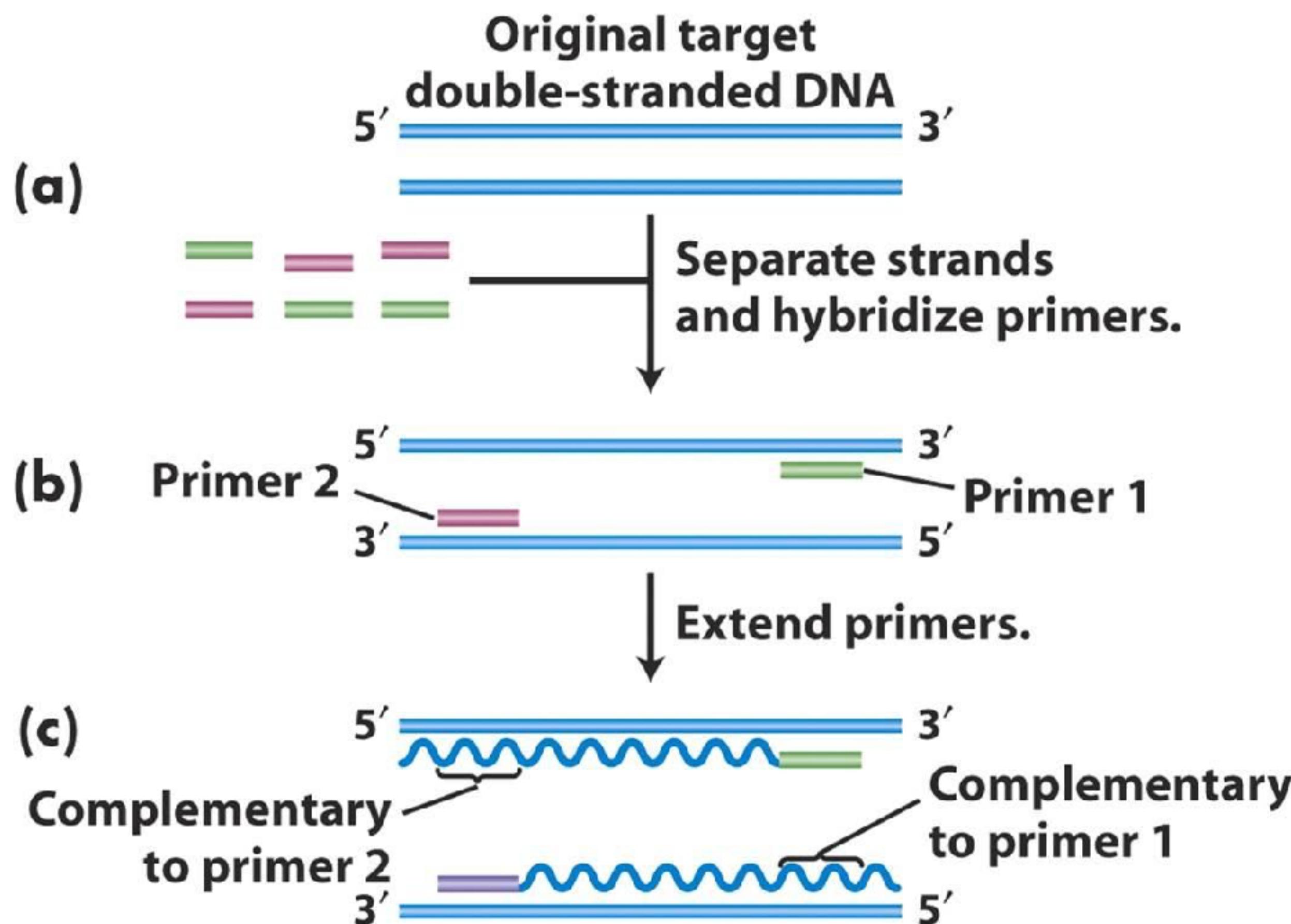


Pyrimidine
Deoxyribonucleotides

DNA SEQUENCING

DNA EXTRACTION AND PREPARATION

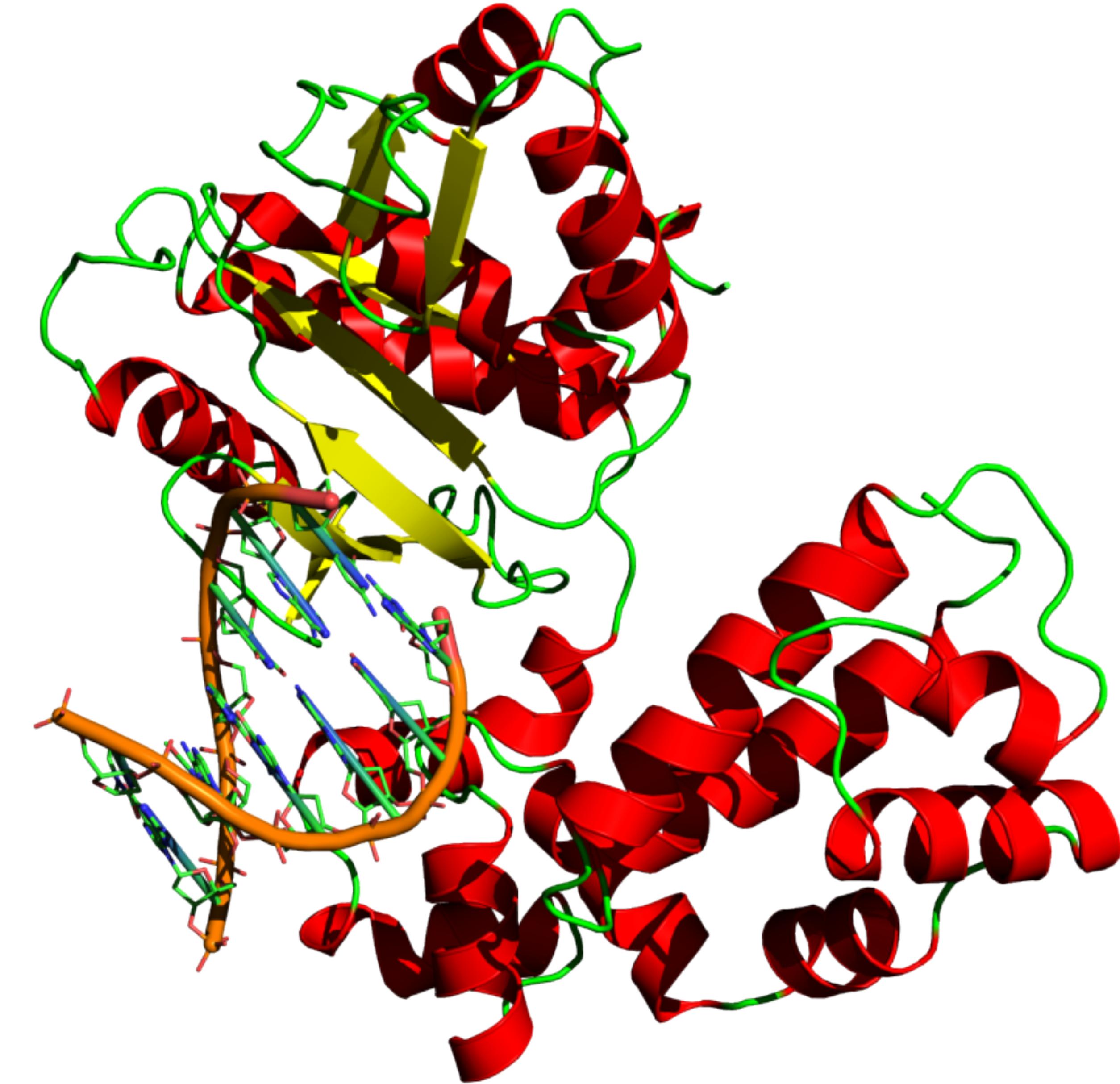
- Primers
 - "Designer" strand of nucleic acid
 - Serves as starting point for DNA replication



DNA SEQUENCING

IMPORTANT MOLECULES FOR SEQUENCING

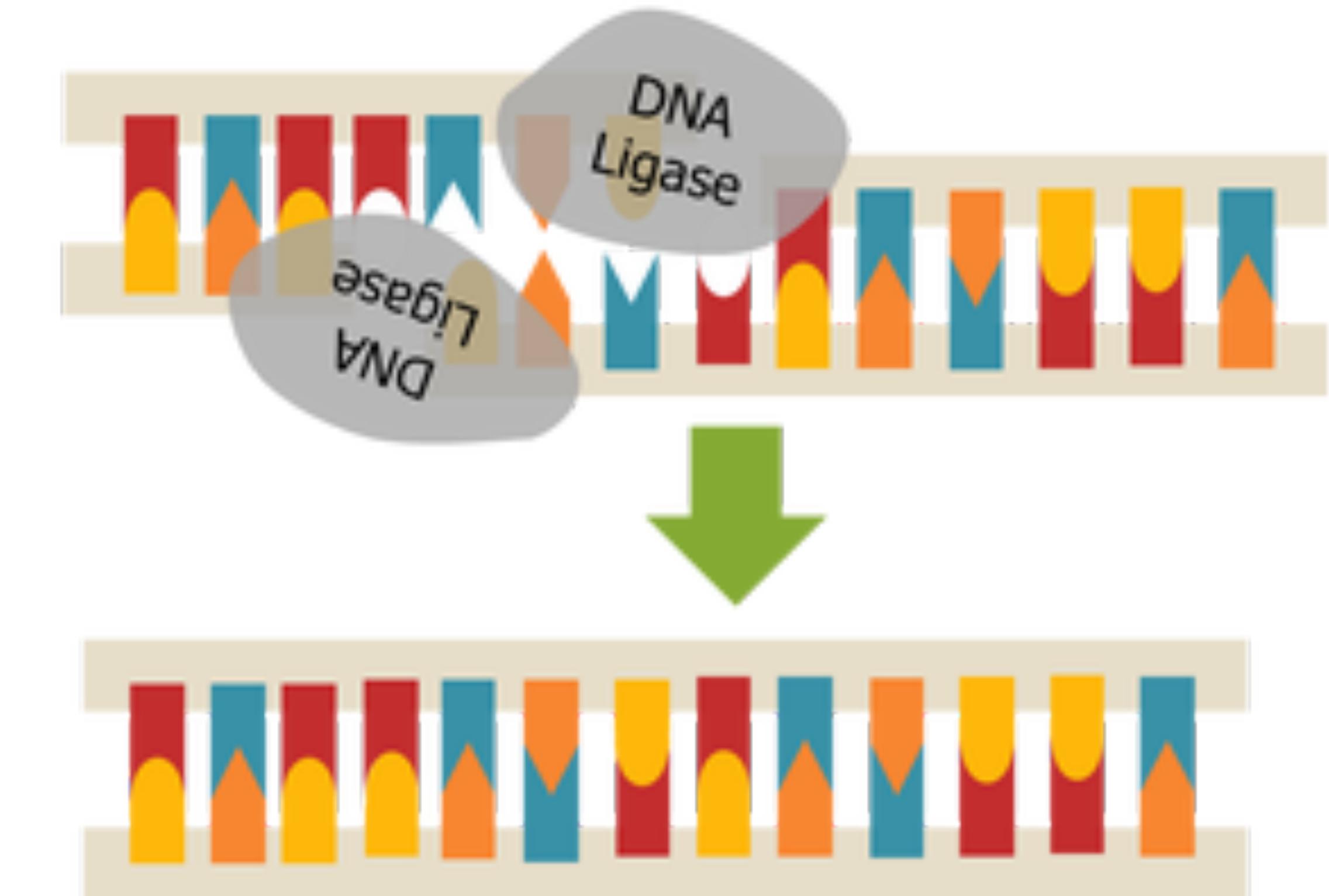
- DNA Polymerase
 - DNA polymerase can add free nucleotides
 - Fills in gaps in DNA (No known DNA polymerase is able to begin a new chain)



DNA SEQUENCING

IMPORTANT MOLECULES FOR SEQUENCING

- Ligase
 - Links together DNA fragments



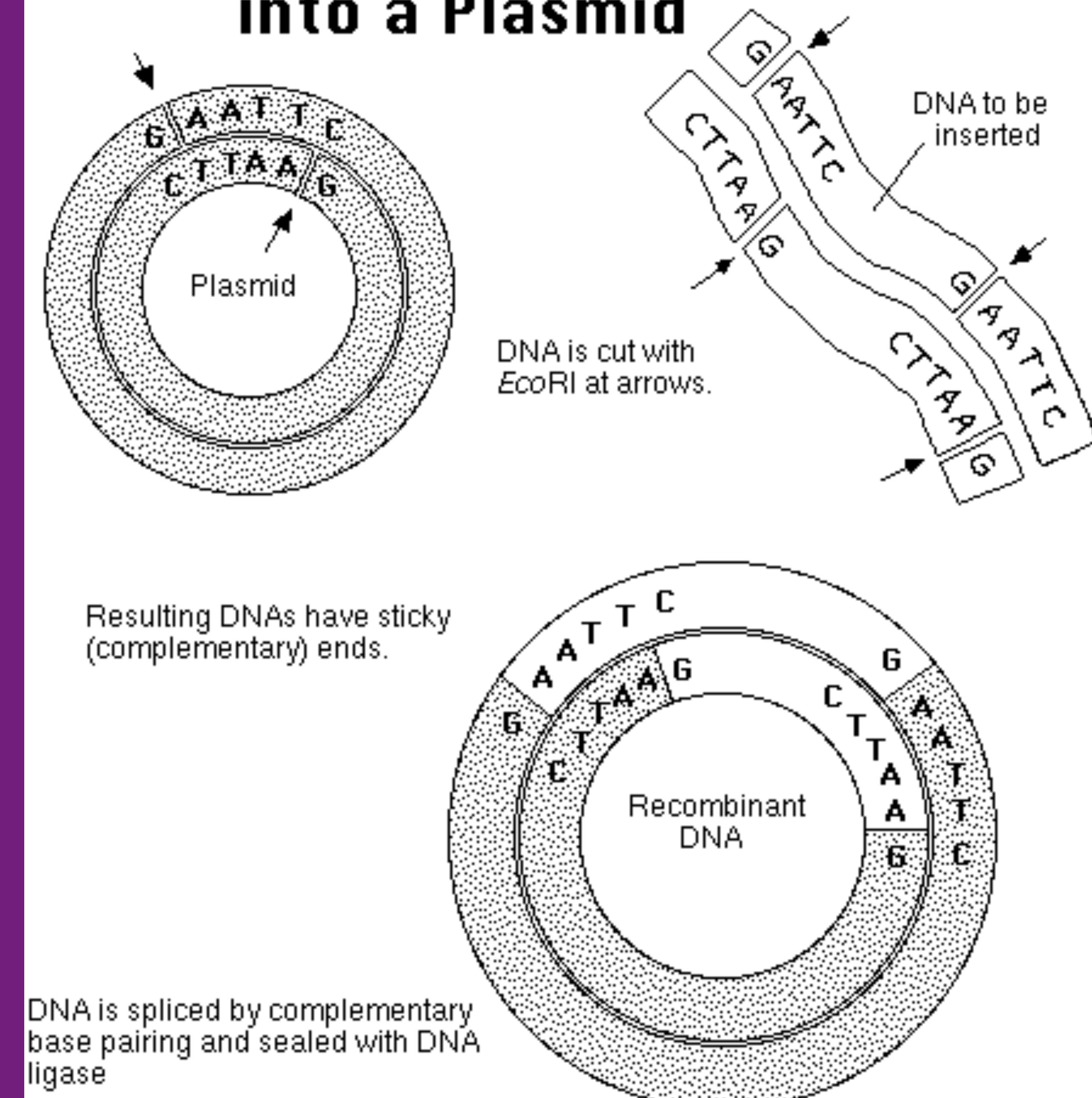
"DNA Ligase" by HeatherTsai Licensed under CC BY-SA 3.0

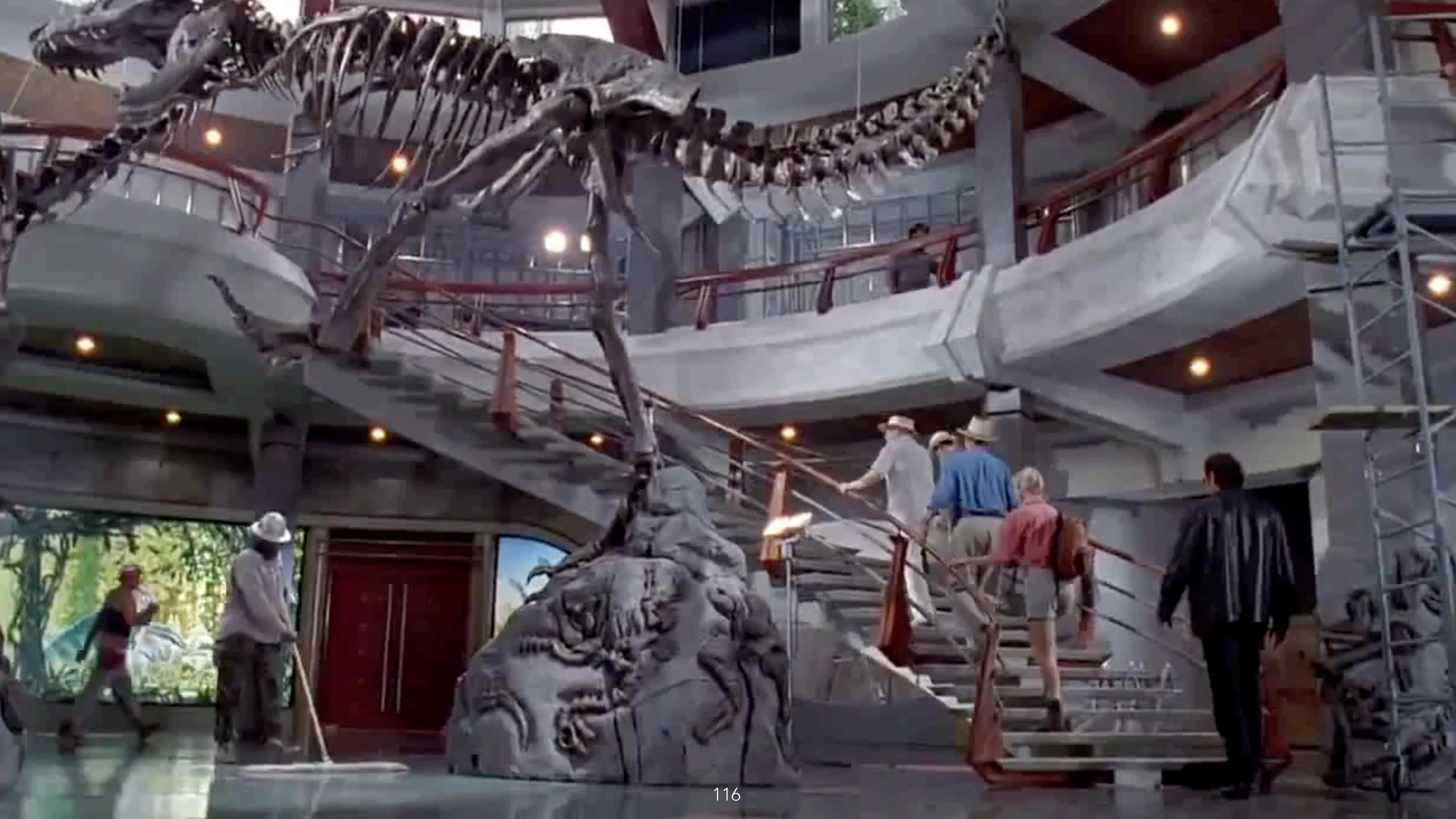
DNA SEQUENCING

IMPORTANT MOLECULES FOR SEQUENCING

- Plasmid
 - Small circular molecules of extrachromosomal, double-stranded DNA
 - Naturally occur both bacteria and yeast
 - They replicate as independent units

Inserting a DNA Sample into a Plasmid





STEP 1: DNA

EXTRACTION

DNA SEQUENCING

DNA EXTRACTION AND PREPARATION

- Break open cells
 - grinding
 - lyse
- Remove cellular proteins and lipids
 - Detergent
- Precipitate DNA
 - Alcohol (DNA is insoluble)



STEP 2: DNA

CLONING

DNA CLONING

- We need a lot of DNA to work with
- Generate many billions of identical DNA molecules
- Approaches
 - Bacterial Cloning
 - PCR
 - cDNA libraries

DNA CLONING

BACTERIAL CLONING

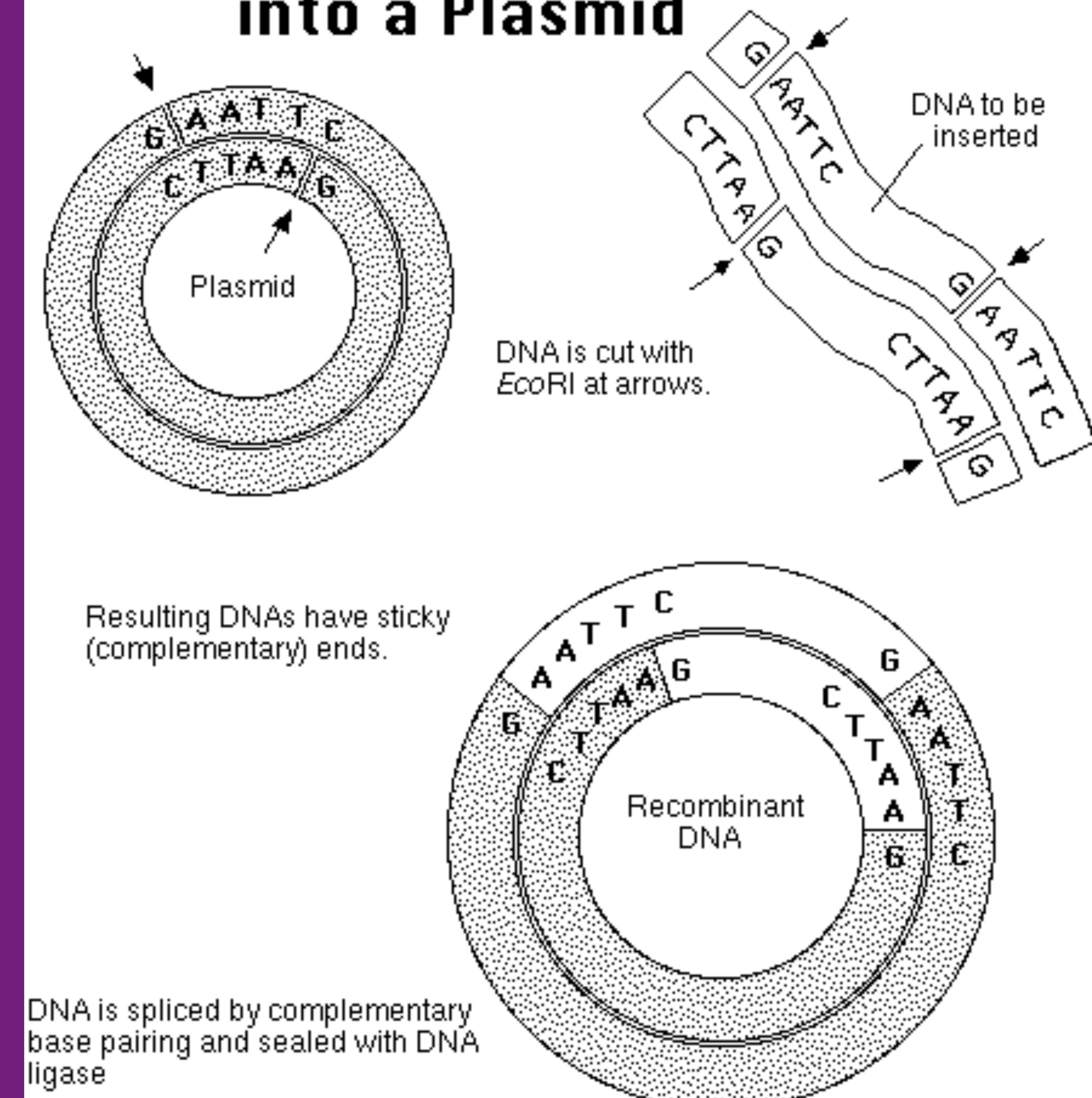


DNA SEQUENCING

BACTERIAL CLONING

- Plasmid
 - Small circular molecules of extrachromosomal, double-stranded DNA
 - Naturally occur both bacteria and yeast
 - Replicate as independent units alongside cell's DNA

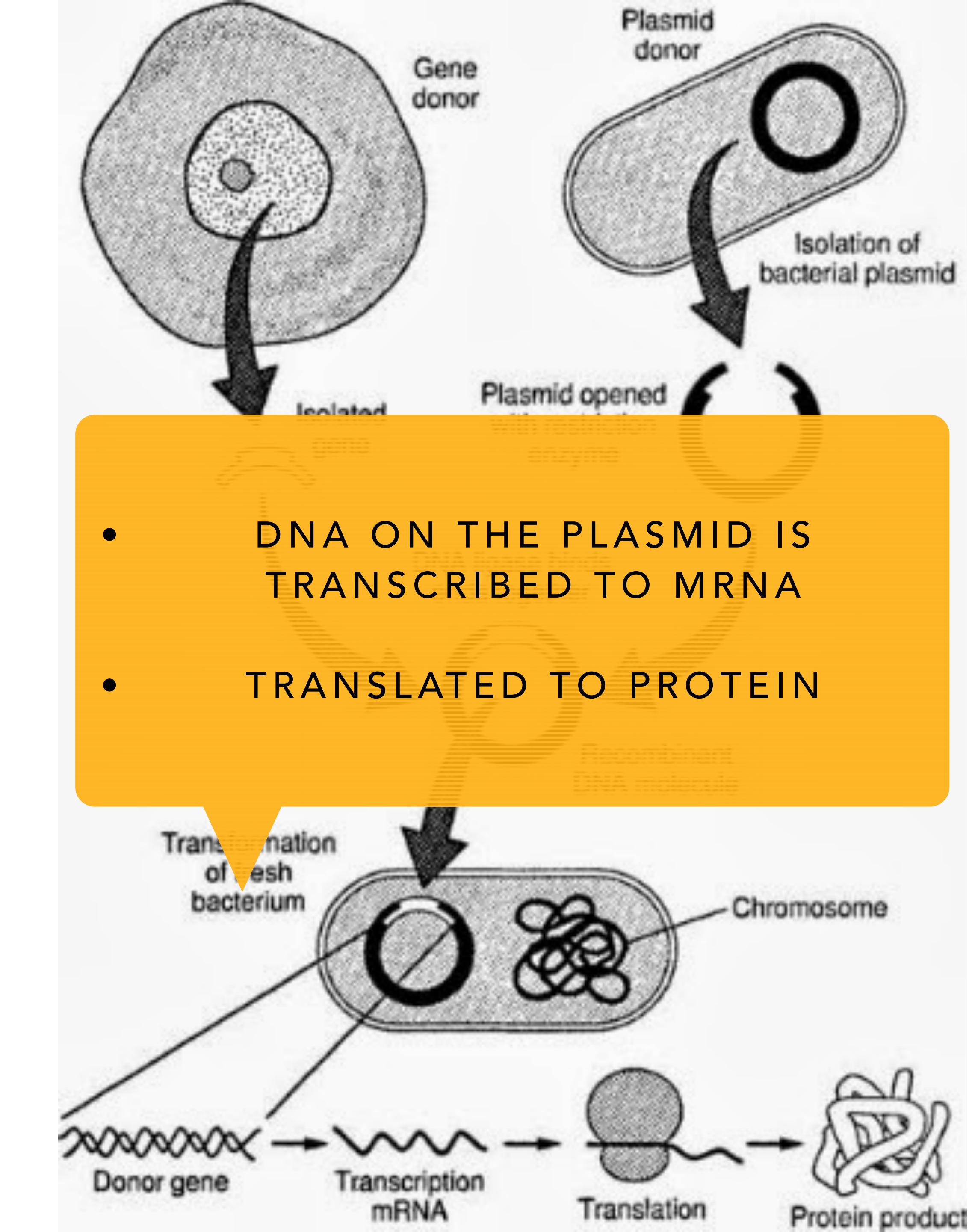
Inserting a DNA Sample into a Plasmid



DNA CLONING

BACTERIAL CLONING

- Bacterial cloning
 - Gene of interest is cut out
 - Cloned into an expression plasmid or vector (transformation)
 - Replicates



DNA CLONING

POLYMERASE CHAIN REACTION

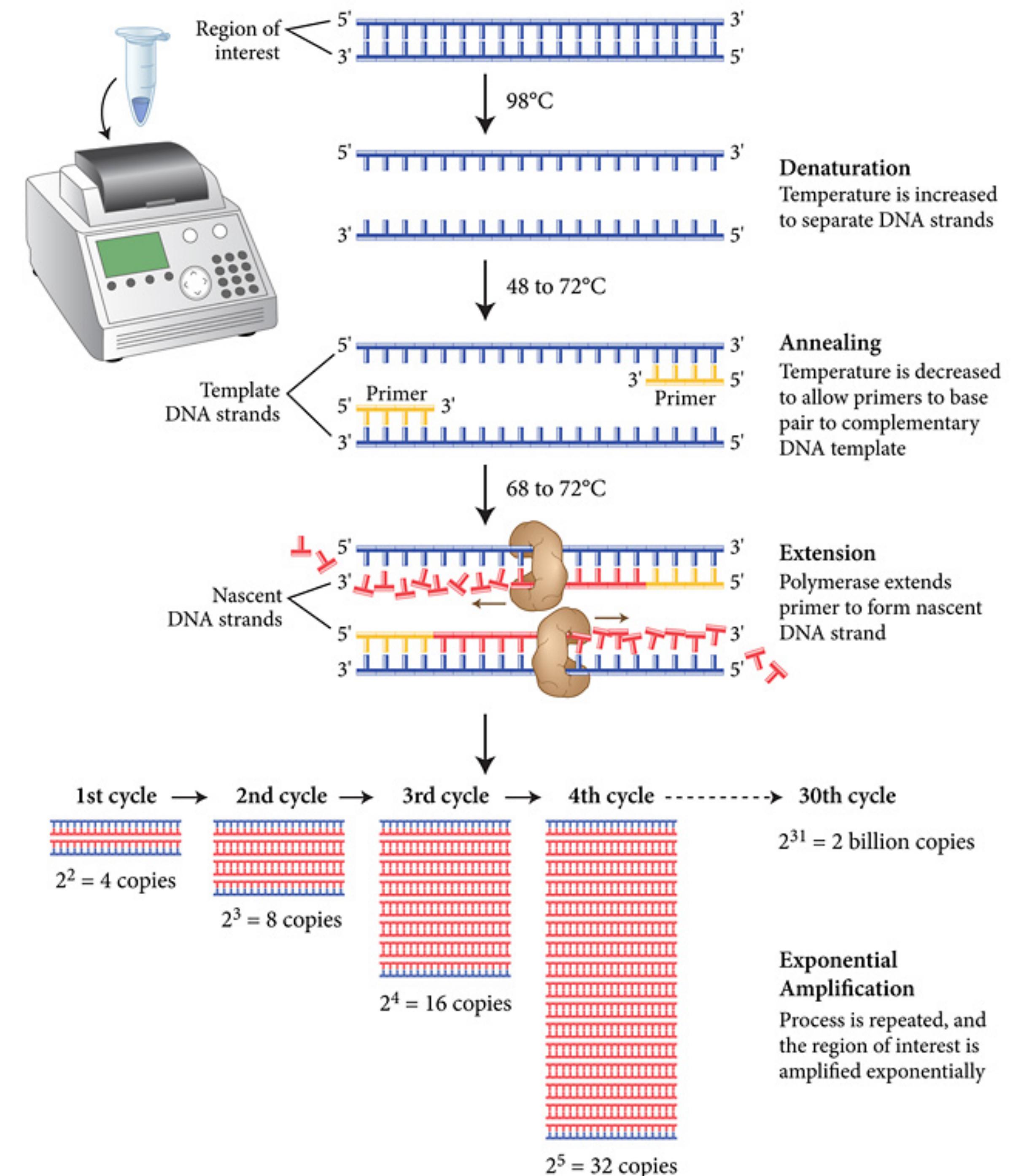


www.dnalc.org

DNA CLONING

POLYMERASE CHAIN REACTION

- Polymerase Chain Reaction (PCR)
 - The DNA sample is mixed in a reaction solution
 - Buffers, free nucleotides, and single-stranded DNA primers
 - DNA polymerase added
 - The reaction mix plus polymerase is then placed in a thermal cycle
 - Each copy can act as a template for subsequent cycles
 - Grows exponentially with each cycle



DNA CLONING

BACTERIAL CLONING VS PCR

- PCR - The Bad:
 - If the DNA target is unknown, there is no way to design specific primers
 - Polymerases have poor fidelity (introduce mutations as they copy)
 - Cannot efficiently produce fragments over a couple kb in length
- PCR- The Good:
 - Good at amplifying tiny amounts of DNA

DNA CLONING

BACTERIAL CLONING VS PCR

- Bacterial Cloning: The Bad
 - Several of the steps to cloning a fragment have relatively low efficiencies
 - Require a relatively large amount of starting material for success
- Bacterial Cloning: The Good
 - Produces larger (miligram) quantities of DNA
 - Can provide clues into gene function (via cloning artifacts)

DNA CLONING

Gene cloning vs PCR

Gene cloning	PCR
A gene amplification technique where Recombinant DNA is constructed invitro and is amplified invivo inside a bacterium.	A gene amplification technique where the DNA is amplified in vitro. No need for the construction of rDNA .
Restriction enzymes, DNA ligase, vector DNA and bacterial cells are required.	Taq DNA polymerase or a thermostable DNA polymerase, RNA primers and free deoxyribonucleotides are required along with DNA segment to be amplified.
At least a microgram quantity of DNA is required for amplification.	A nanogram of DNA is enough for amplification.
A restriction enzyme is required for reisolation of the amplified DNA from rDNA.	No need for reisolation or use of enzymes.
For getting desired DNA, amplified DNA should be screened at the final step.	No need for screening after PCR, if the DNA is pure before starting the reaction.
2-4 days should be spent for an experiment.	Maximum 4 hours is enough for an experiment.
No need for automation.	Automation is a must.
Labour intensive.	No need for intensive labours.
Error possibility is more.	Error possibility is less.
The amplified DNA is put into limited number of uses.	The amplified DNA is put into many uses because of less error possibility.



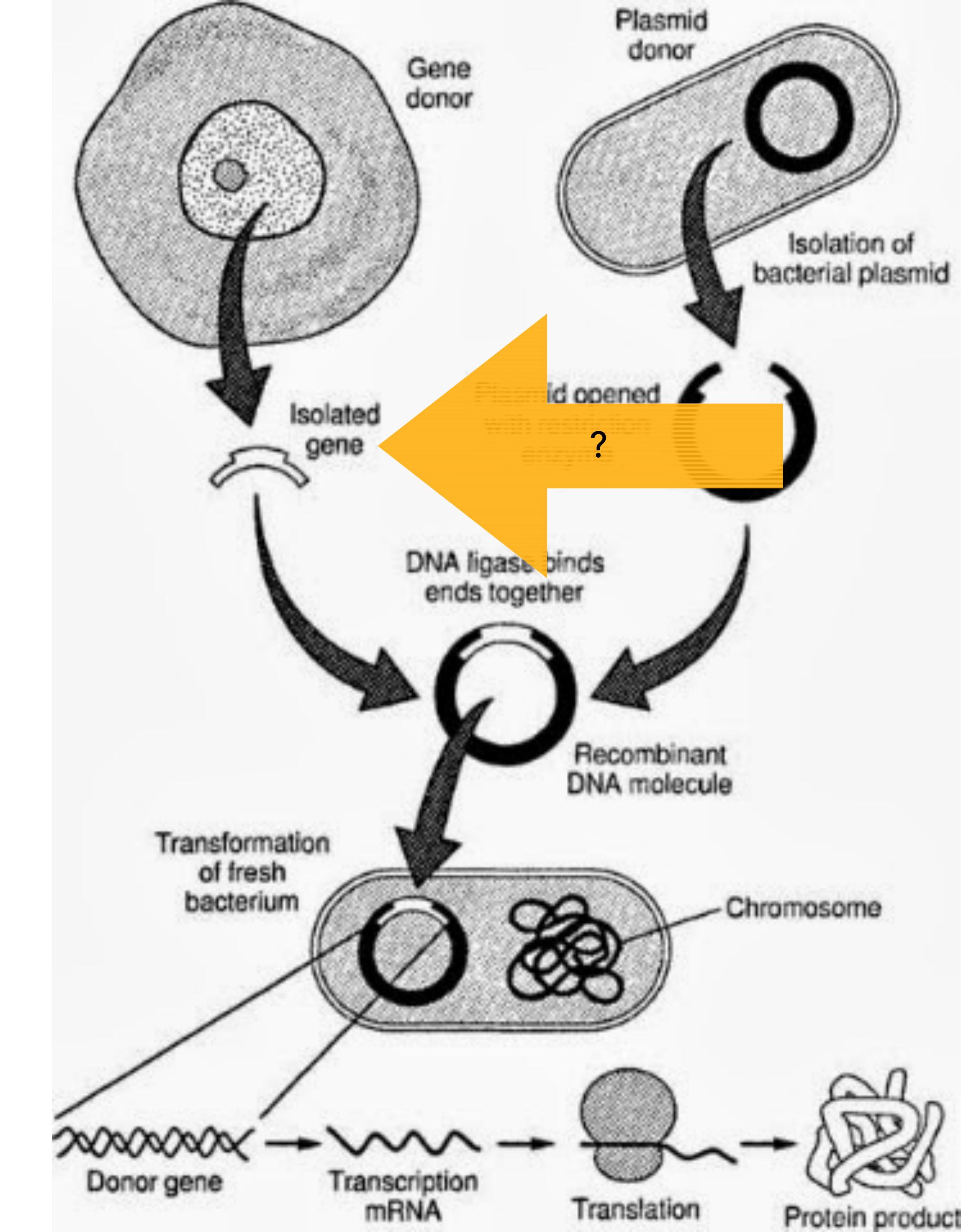
DNA CLONING

- In practice
 - Complimentary techniques
 - PCR used to generate enough samples for bacterial cloning
 - PCR used afters bacterial cloning to isolate a gene

ALT. STEP 2: cDNA LIBRARIES

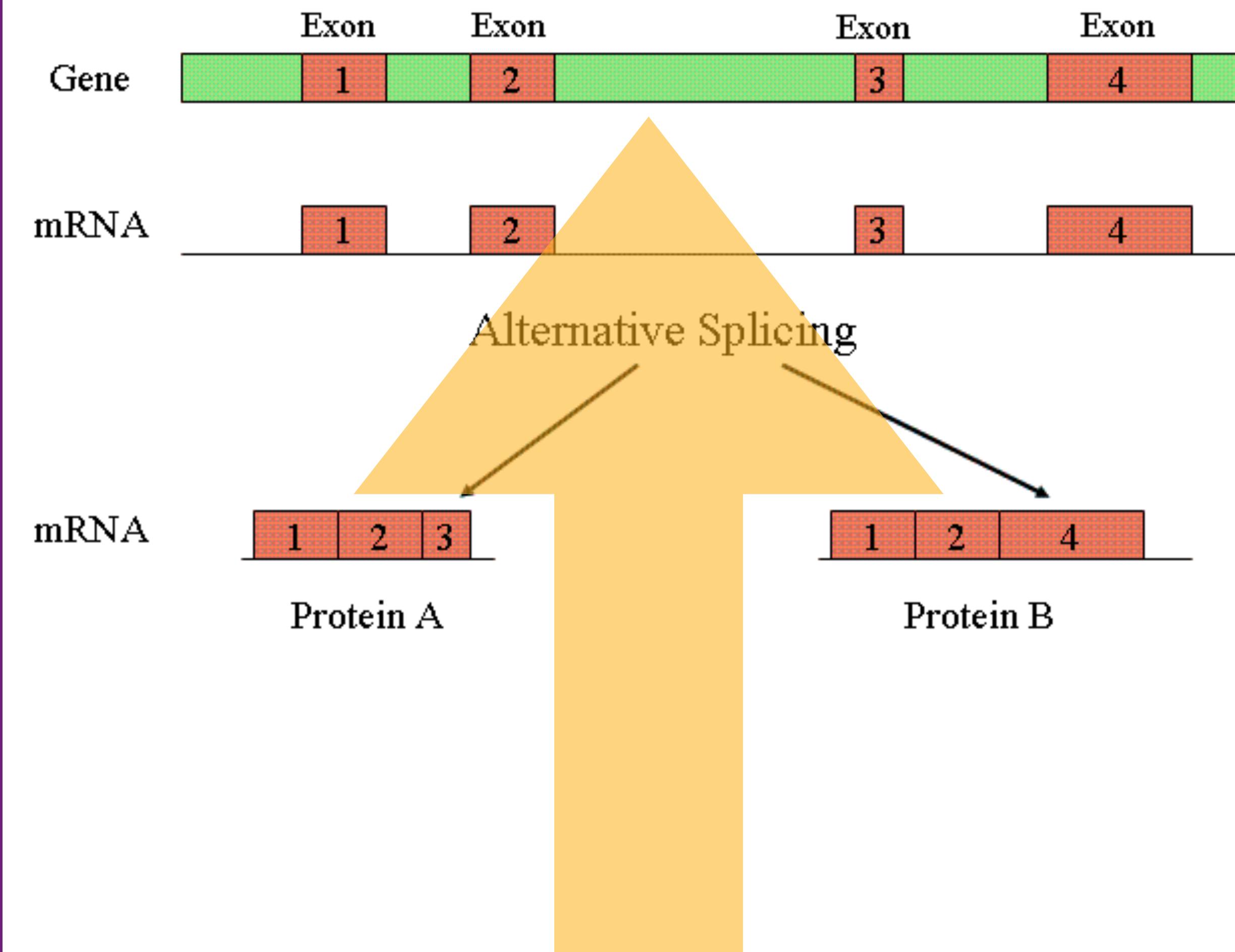
CDNA LIBRARIES

- Cloning assumes we know what DNA/gene we are looking for
- What if you don't know the sequence of the gene you are interested in?
 - Introns/exons
 - Alternate splicing

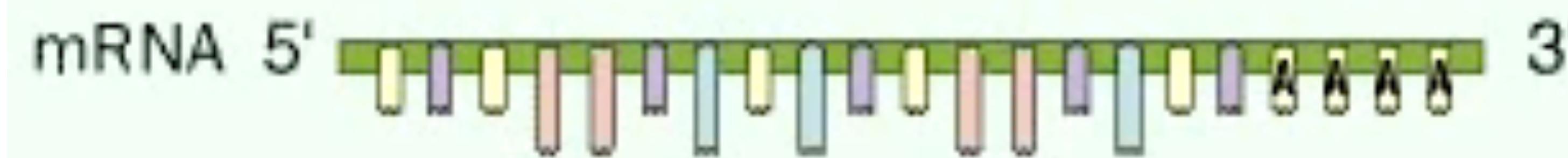


CDNA LIBRARIES

- Use mRNA to find out the DNA it was transcribed from
 - Alternative way to get DNA of expressed genes

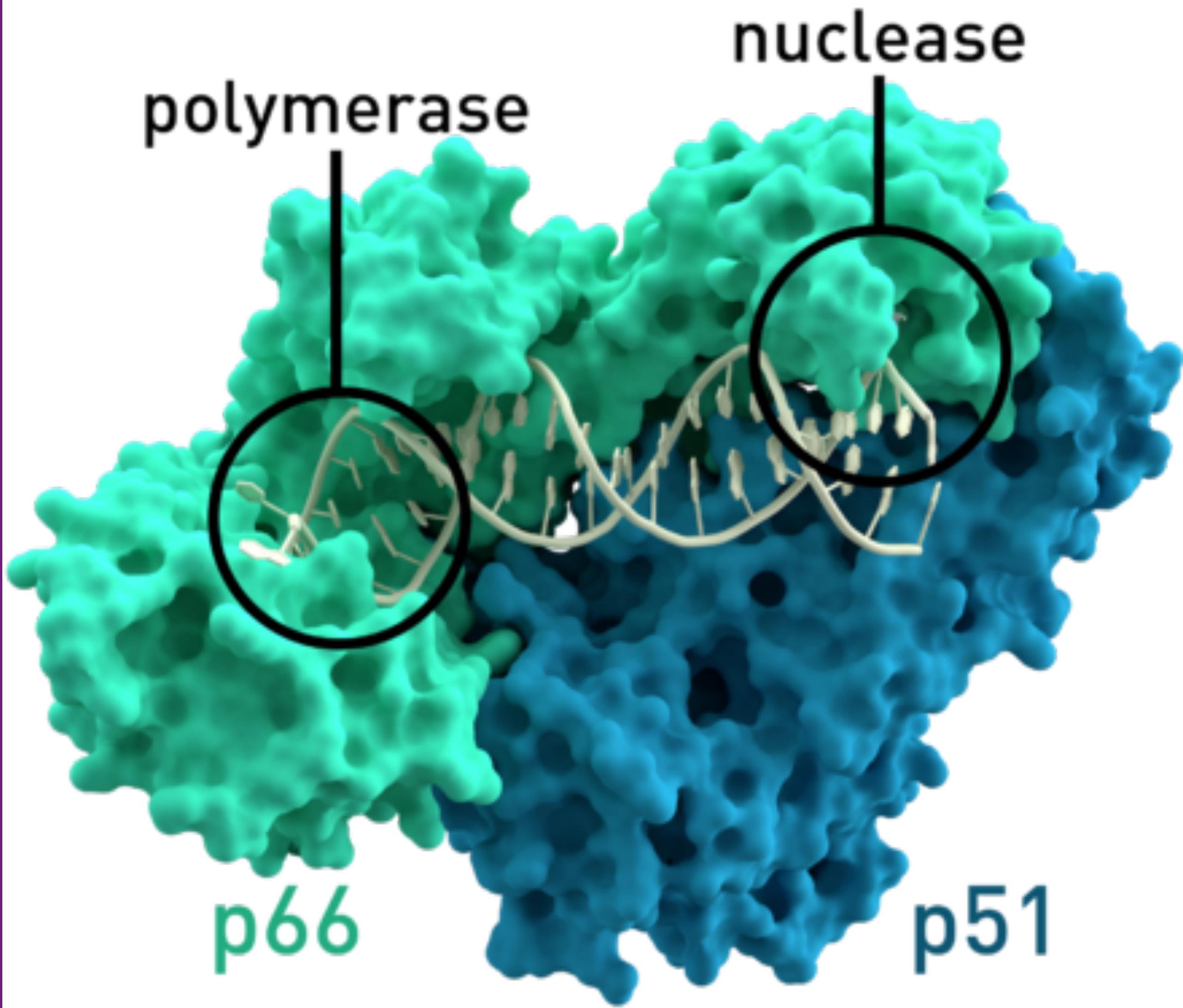


CDNA LIBRARIES



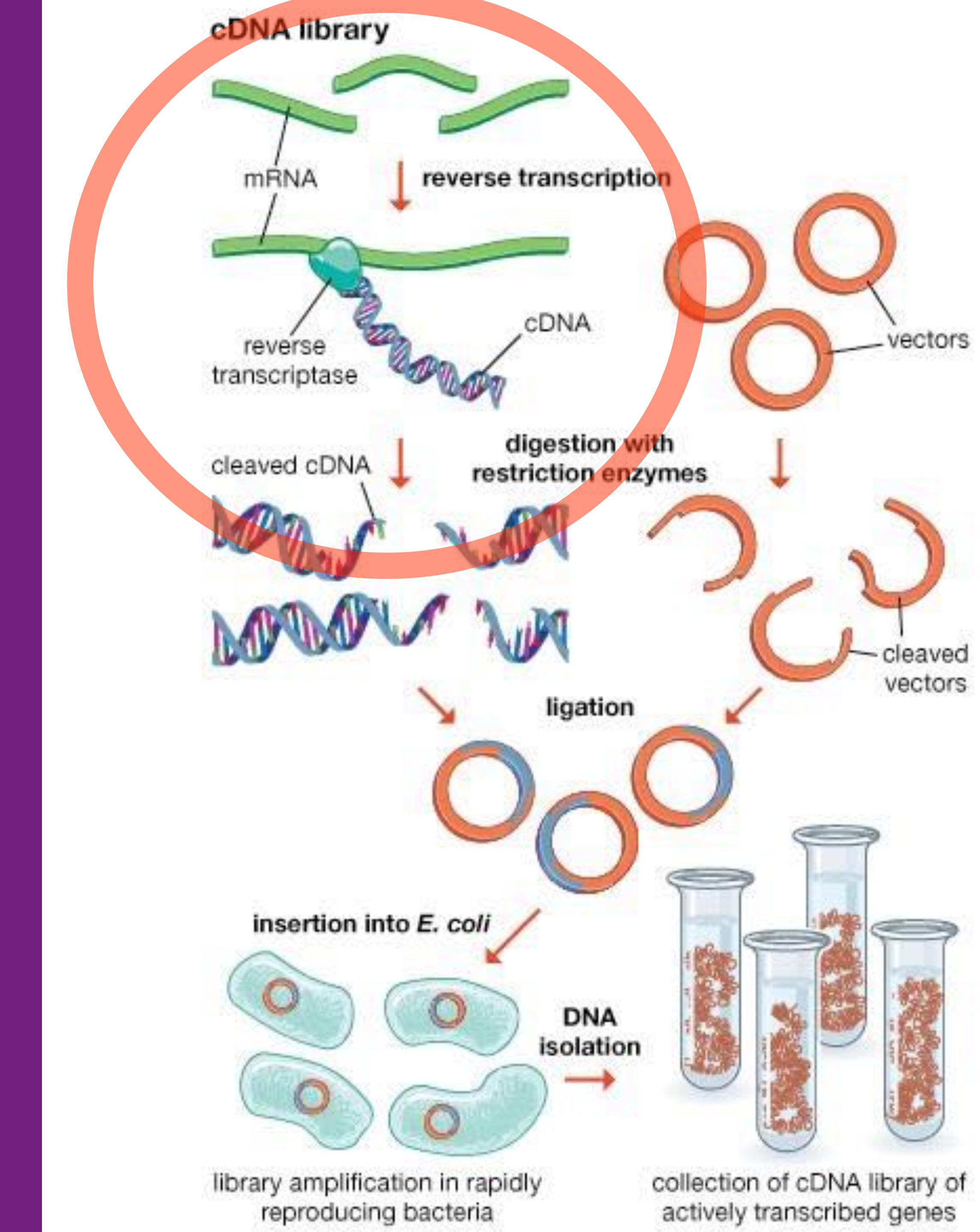
CDNA LIBRARIES

- Reverse transcriptase
 - Enzyme that generates cDNA from RNA template
 - RNA-dependent DNA polymerase activity
 - Used by some viruses



CDNA LIBRARIES

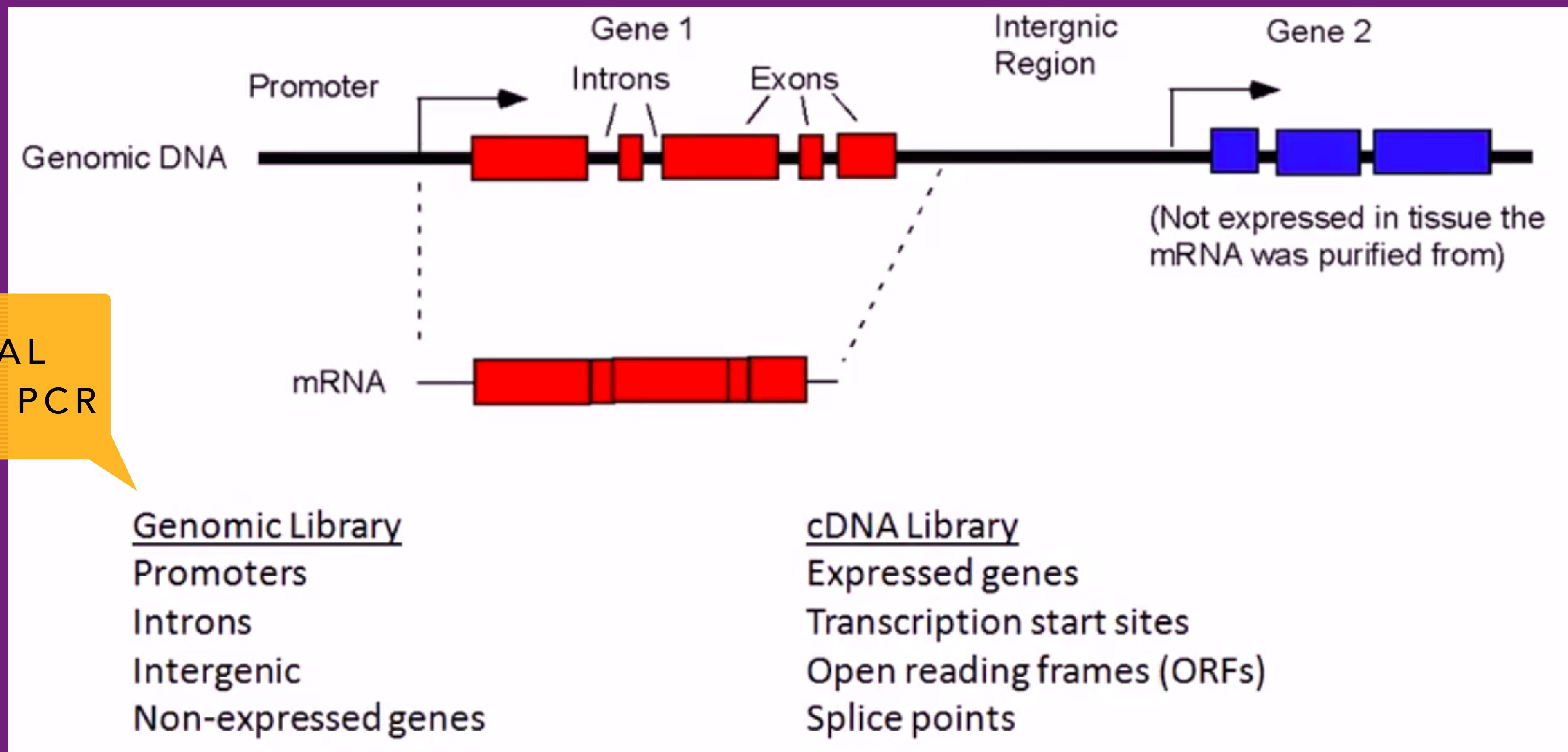
- Cloned complimentary DNA (cDNA) libraries
 - Produced from fully transcribed RNA
 - Contains expressed genes
 - Introns already spliced out
 - Can be expressed in bacterial cell
 - Lacks regulatory elements found in genomic DNA library



CDNA LIBRARIES

- RNA is relatively unstable to work with
- It is typically converted to complimentary DNA (cDNA)
 - Many databases contain cDNA sequences corresponding to RNA transcripts
- Database naming conventions
 - Term cDNA is also used to refer to an mRNA transcript's sequence, expressed as DNA bases (GCAT) rather than RNA bases (GCAU)

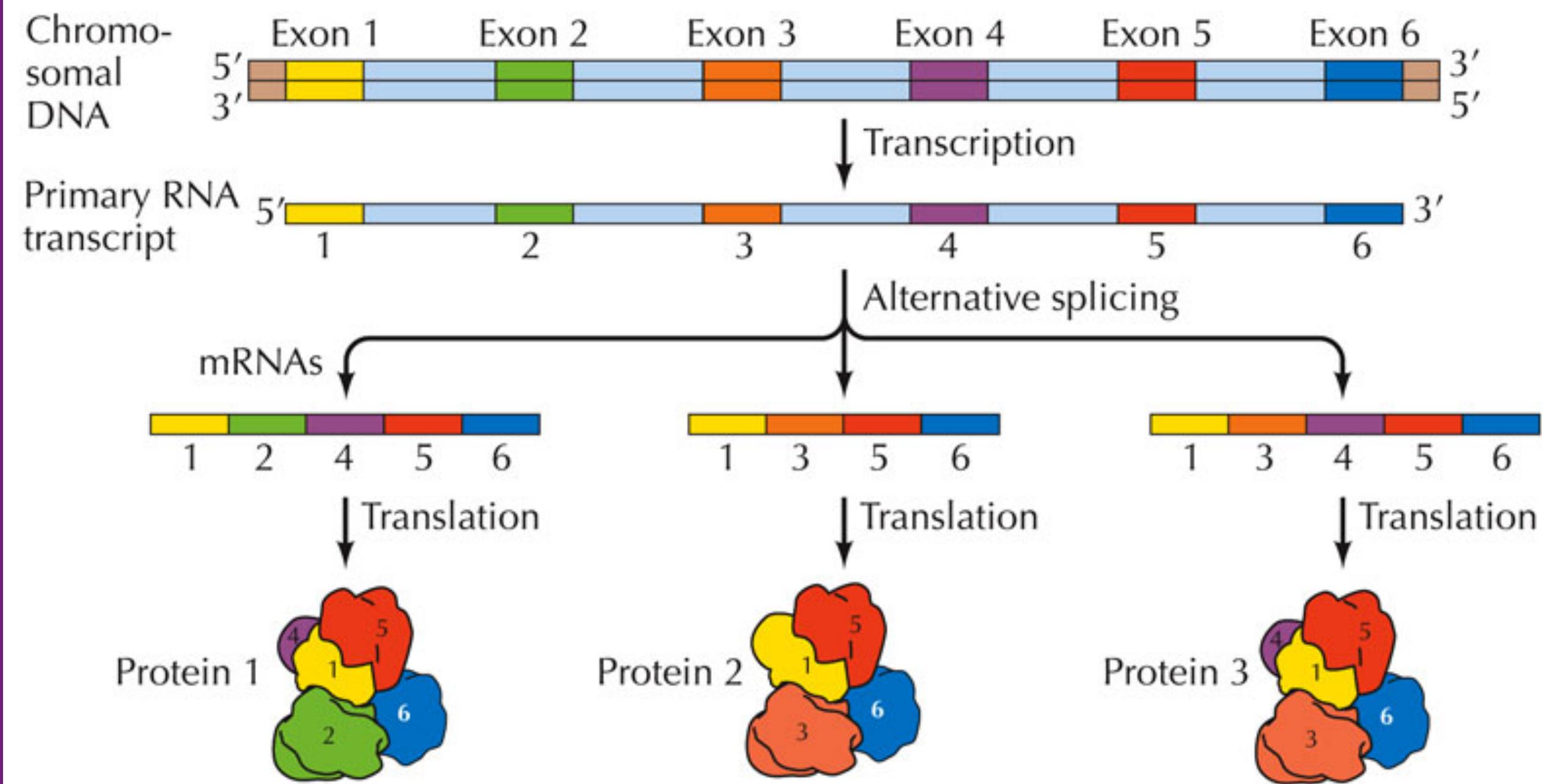
CDNA LIBRARIES



- Differences between genomic and cDNA library

CDNA LIBRARIES

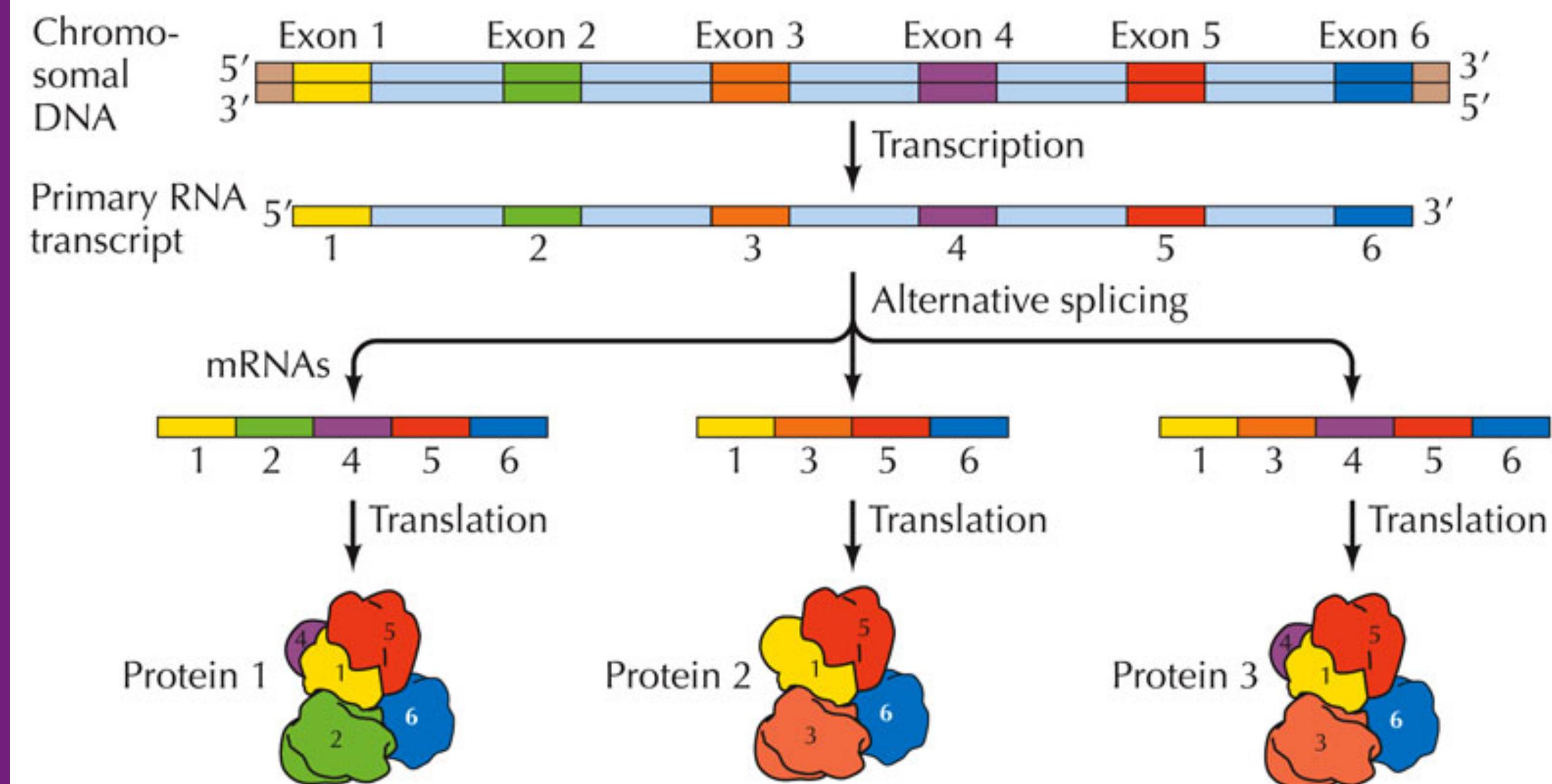
- Alternative splicing
 - Different mRNAs from same premRNA transcript
 - Exons shuffling allows single gene to code for multiple proteins
 - Regulated process during gene expression



THE CELL, Fourth Edition, Figure 5.5 © 2006 ASM Press and Sinauer Associates, Inc.

CDNA LIBRARIES

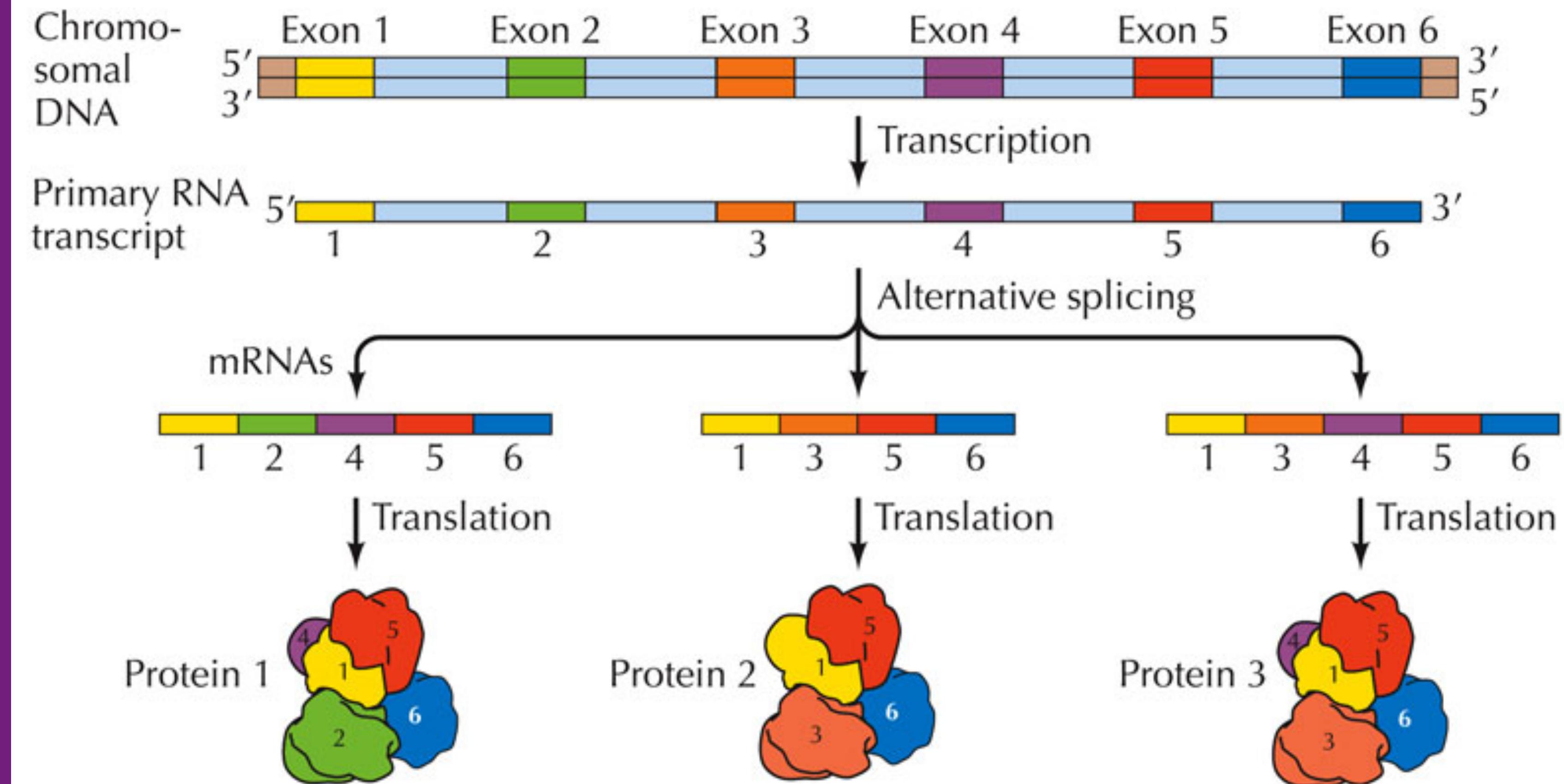
- ~80% of human genome are alternatively spliced
- On avg. 8 different alternatively splices per gene
 - ~22,000 protein coding genes
 - $*8 = 176,000$ proteins
 - Human Genome initial estimates were ~150,000 genes
 - Did not take into account alternate splicing



IN DATABASES (E.G. NCBI REFSEQ) YOU WILL FIND MANY MRNAS "SPLICE VARIANTS"

CDNA LIBRARIES

- cDNA library from a gene does not guarantee you can easily discover the DNA sequence



BIOINFORMATICS

(FOR COMPUTER SCIENTISTS)

MPCS56420
SPRING 2020
SESSION 2

