

**Biotechnology  
in  
Molecular Biology**

**Dr Piyush B. Tailor**

Associate Professor

Department of Biochemistry

Govt. Medical College, Surat

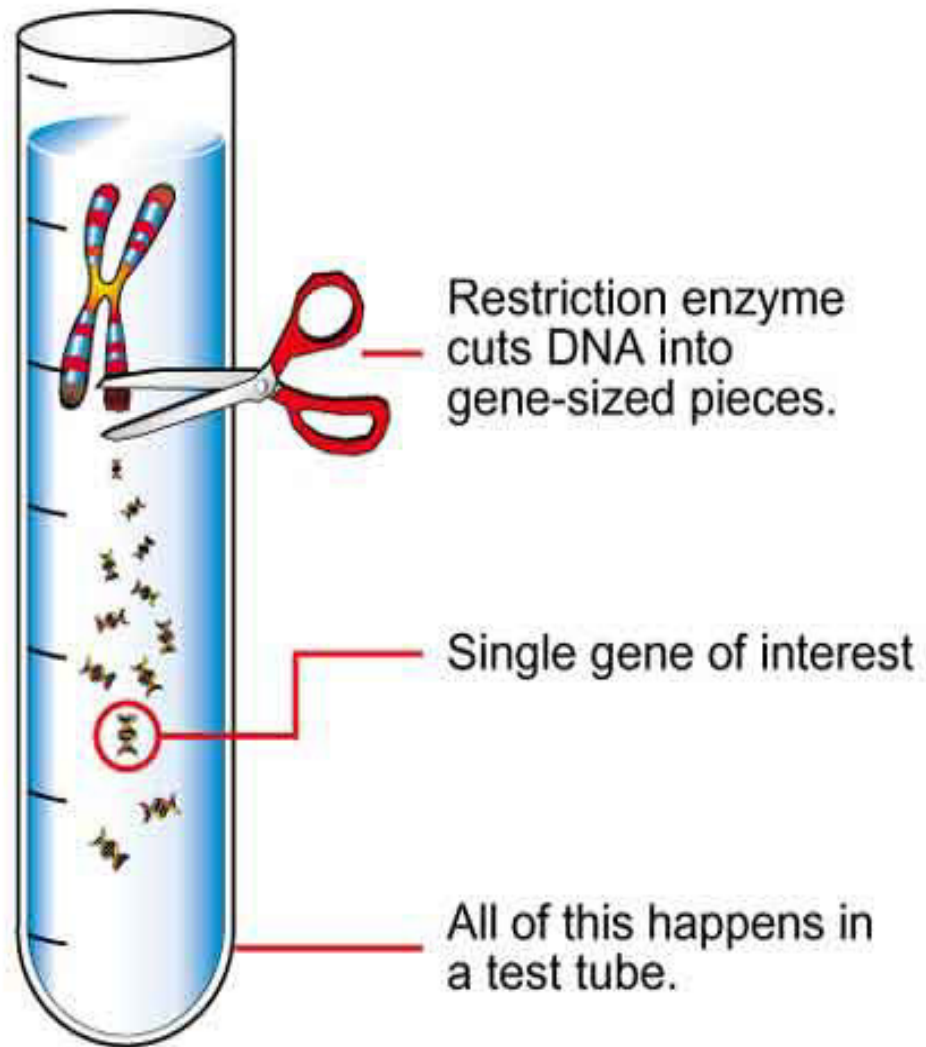
# What is Biotechnology?

**With ( $10^9$ ) DNA Base Pair & (20,000 to 30,000) Gene**

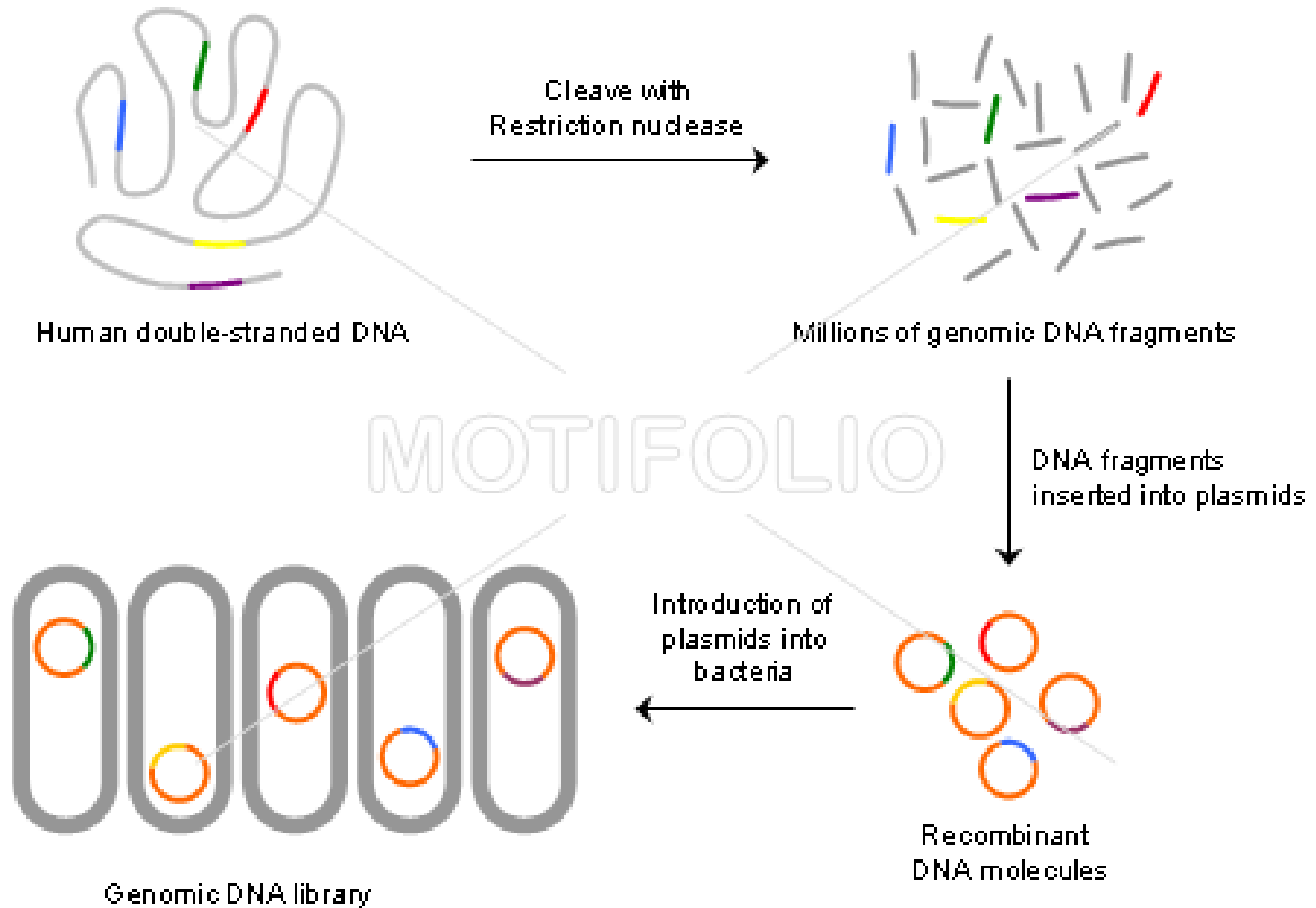
- Restricted Endonuclease Enzyme
- DNA Fragement
- Identification of Gene
- Human Genomic Project
- Gene Library
- **Use of All of Above for Clinical – Diagnosis and Management**

Dr Piyush Tailor

# Molecular Biotechnology



# Construction of a human genomic DNA library

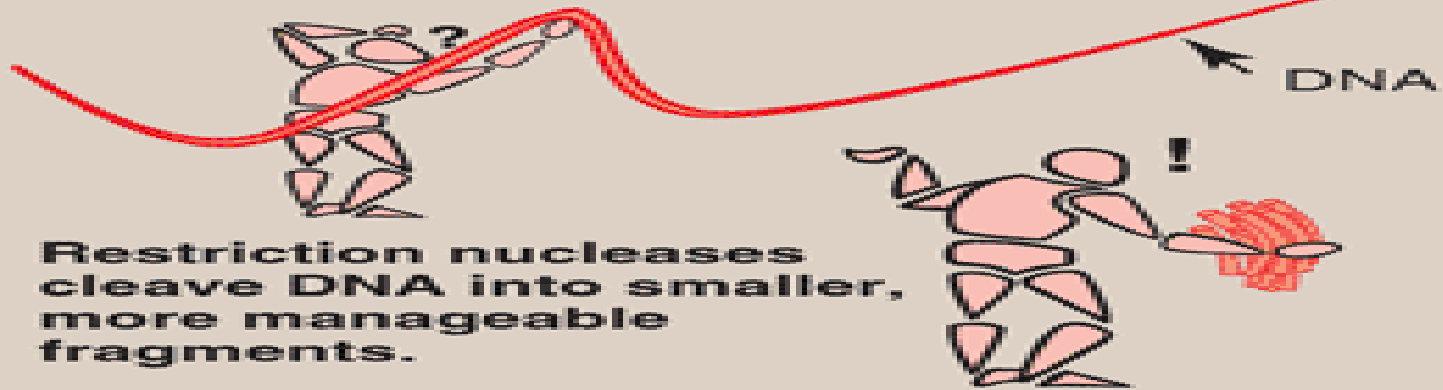


MOTIFOLIO

# Goal of Human Genomic Project

- Identify all of the genes in human DNA.
- Determine the sequence of the 3 billion nucleotide bases of DNA
- Store this information in Data bases.
- To Know function of DNA gene
- Develop efficient sequencing technologies.
- Develop tools for data analysis.

## Restriction endonucleases



Restriction nucleases cleave DNA into smaller, more manageable fragments.

## Cloning of DNA



DNA fragments must be amplified to be more useful.

## Probes



A specific fragment can be identified using a complementary probe.

# Restricted Endonuclease

- Why called “Restriction” Enzymes
  - Restrict Growth of Bacteriophages
- Each cleaves at a **specific nucleotide sequence**
- Used to obtain defined DNA segments
  - **Restriction Fragment.**
- Cleave double-stranded (ds) DNA
- Easy for DNA analysis.

Dr. Piyush Tailor

# Action of Restricted Endonuclease

A

ATCGATCGATCGATCGA|AATTCATCGATCGCTAG  
TAGCTAGCTAGCTAGCTTAAG|GTAGCTAGCGATC

B

ATCGATCGATCGATCGAATTCATCGATCGCTAG  
TAGCTAGCTAGCTAGCTTAAGGTAGCTAGCGATC

C

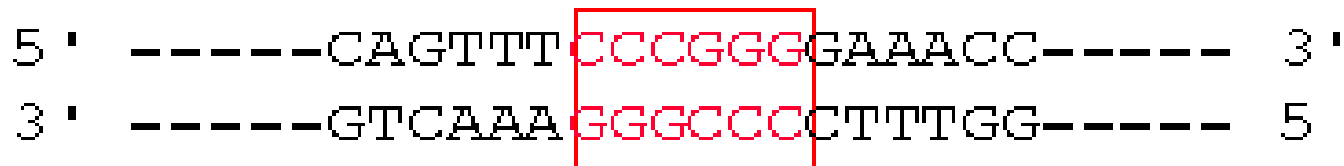
ATCGATCGATCGATCG                      AATTCATCGATCGCTAG  
TAGCTAGCTAGCTAGCTTAA                      GTAGCTAGCGATC

- Which bond formation is required to rejoin the sequence?
- Which enzyme can make that bonding?



# Action of Restricted Endonuclease

## Sma I



- Which bond formation is require to rejoin the sequence?
- Which enzyme can make that bonding?
- What is advantage or disadvantage of both type of R.E.?

# Specificity of Restriction Endonuclease

- Recognize short specific nucleotide sequences
- Generally **Four or Six base pairs**
- 4 nucleotide pair = cut at every 256 bp
- 6 nucleotide pair = cut at every 4096 bp
- Mostly “**Palindrome Sequence**”

Dr Piyush Tailor

## A Palindrome

When read in the 5' → 3' direction, the sequence on the "top" strand is identical to that of the "bottom" strand.



5' -GAATTC- 3'

3' -CTTAAG- 5'

# “Sticky” and “Blunt” ends

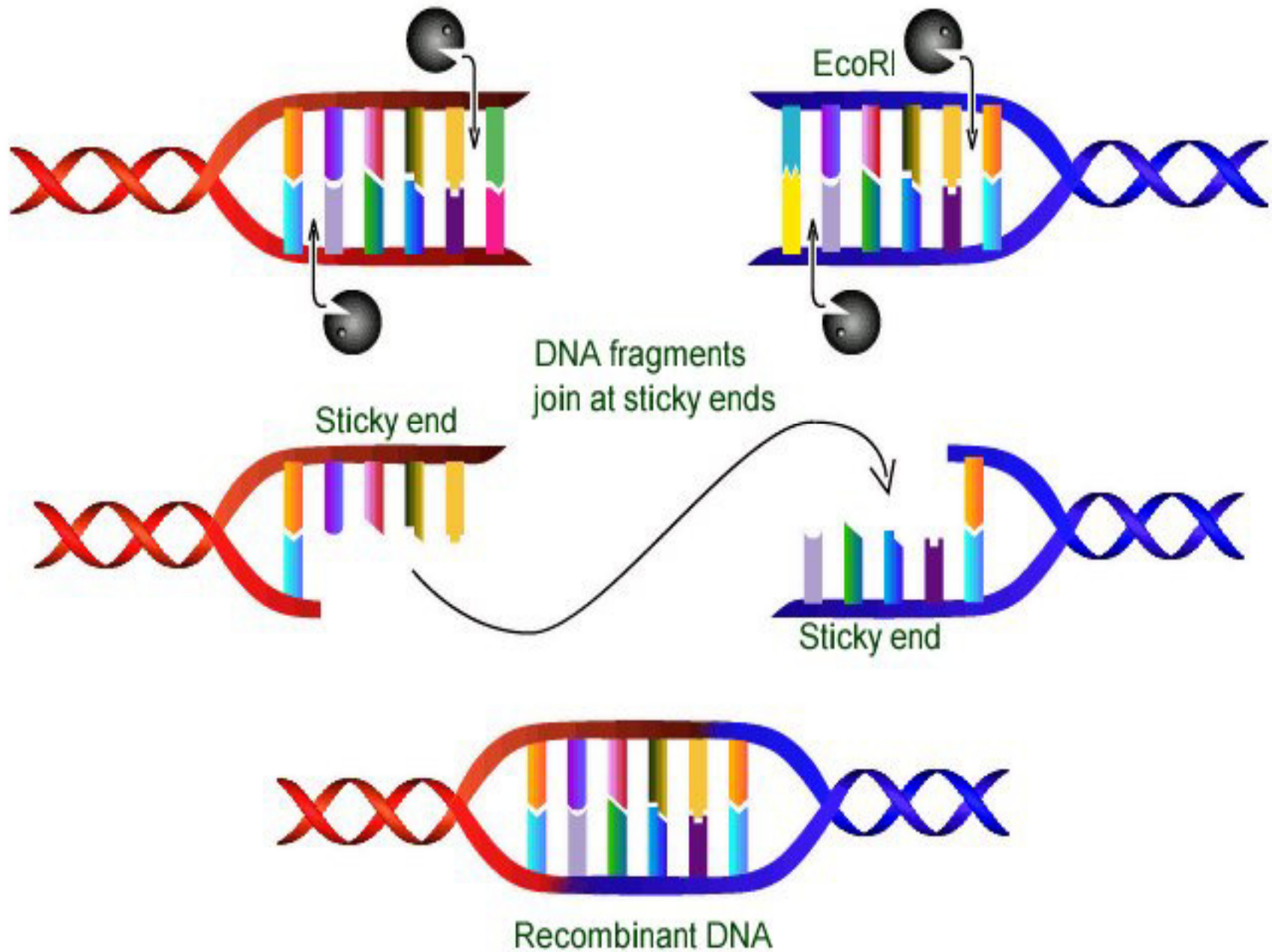
## **STICKY END**

- **TaqI** = produce “sticky” end
- Resulting DNA fragments = single-stranded (ss) sequences
- Sequence is complementary to each other

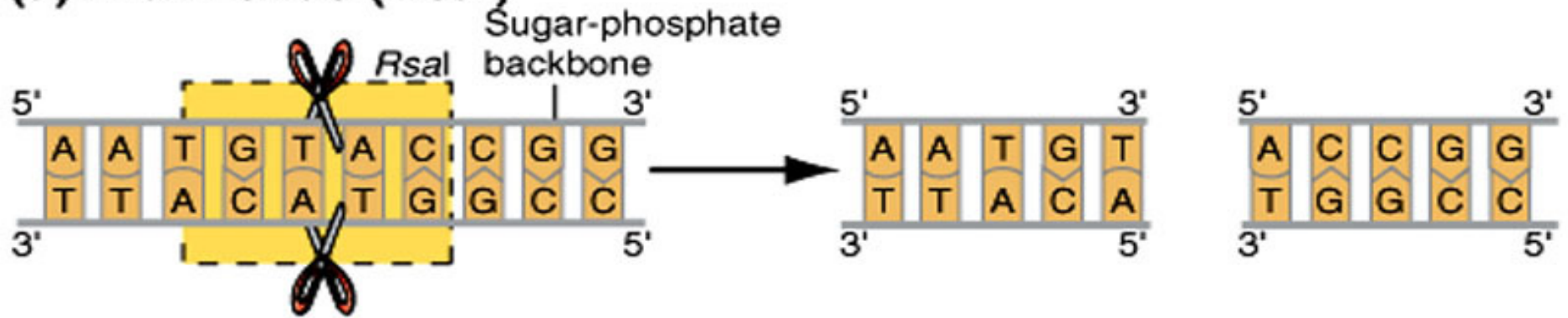
## **BLUNT END**

- **HaeIII** = produce “blunt” ends
- Resulting fragment is double-stranded
- Do not form hydrogen bond
- DNA ligase help to rejoin sticky ends of a DNA fragment

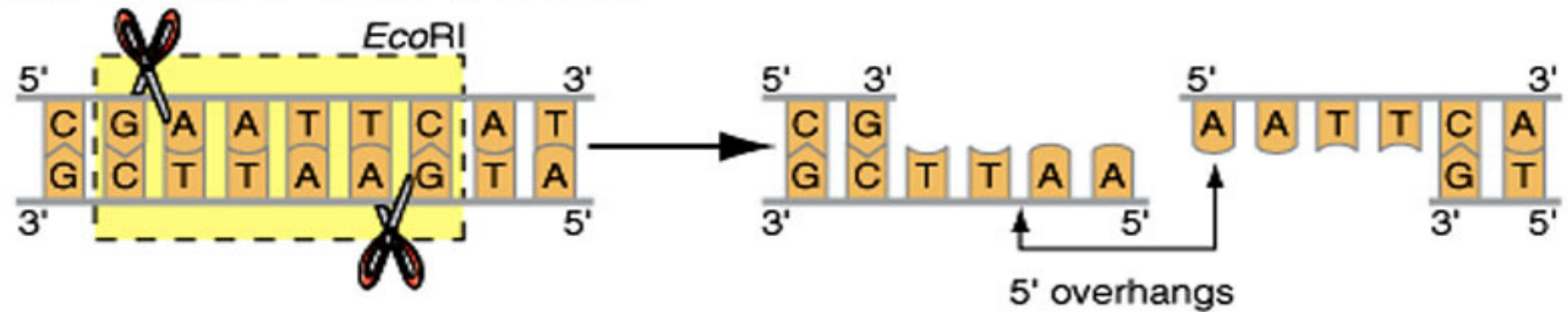
# Restriction Enzyme



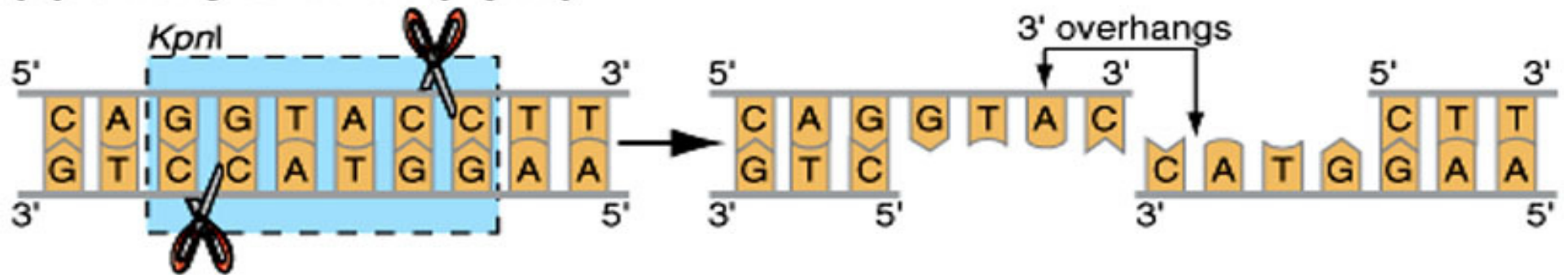
### (a) Blunt ends (*RsaI*)



### (b) Sticky 5' ends (*EcoRI*)



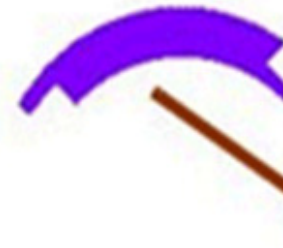
### (c) Sticky 3' ends (*KpnI*)



Foreign DNA  
(Target Gene containing DNA)

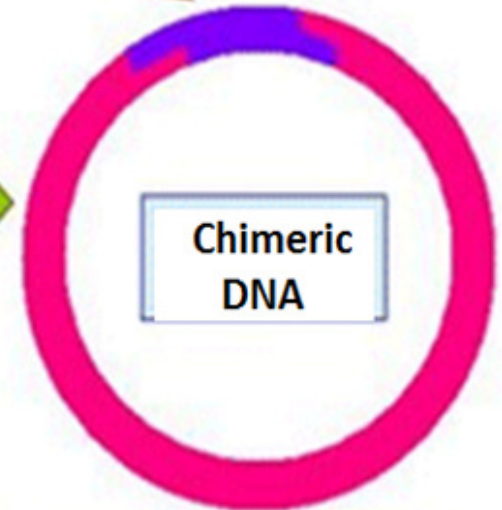
Cutting of Foreign DNA  
through R.E. to extract  
Target Gene

Same  
Restricted Endonuclease  
enzyme for  
Vector & Foreign DNA



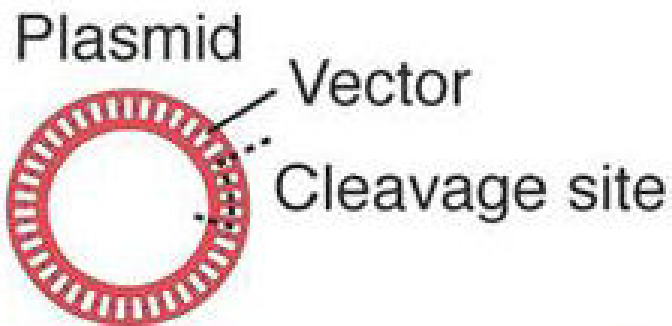
Vector

Plasmid vector has a  
single restricted site,  
hence there is only one  
nick

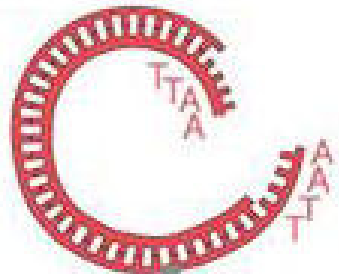


Chimeric  
DNA

Cleaved foreign DNA  
anneales with Plamid  
vestor & Ligated



Cleavage by *EcoRI* endonuclease



Donor DNA

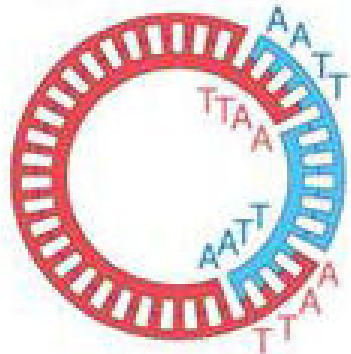


Cleavage sites

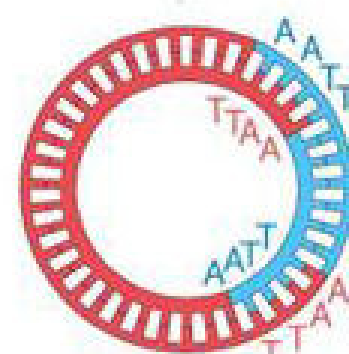
Cleavage by *EcoRI* endonuclease



Hybridization



DNA ligase



Recombinant plasmid



# Restriction Site of DNA

- Sites are recognized by restriction endonuclease
- Cleave DNA into fragments of different sizes

## ➤ Example 1

- Enzyme that recognizes a specific **four-base-pair** sequence
- Produces **Many cuts and Shorter pieces** in DNA.

## ➤ Example 2

- Enzyme requiring sequence of **six base pairs**
- Produces **Fewer cuts & Longer pieces**.
- Different enzyme = Different cleavage

DNA of interest      Second DNA fragment

Cleavage by *TaqI*



DNA of interest with a "sticky" end produced by *TaqI*.



Second fragment of DNA also with a "sticky" end produced by cleavage with *TaqI*.



Cohesive ends of two DNA fragments form hydrogen bonds.

*DNA ligase*



*DNA ligase* forms a 3'→5' phosphodiester bond between the fragments.

## Nomenclature of R.E.

- 1<sup>st</sup> Letter = Genus of bacterium.
- 2<sup>nd</sup> Letters = Name of species.
- Additional letter = Type or Strain
- Final number = order discovered

Dr Piyush Tailor

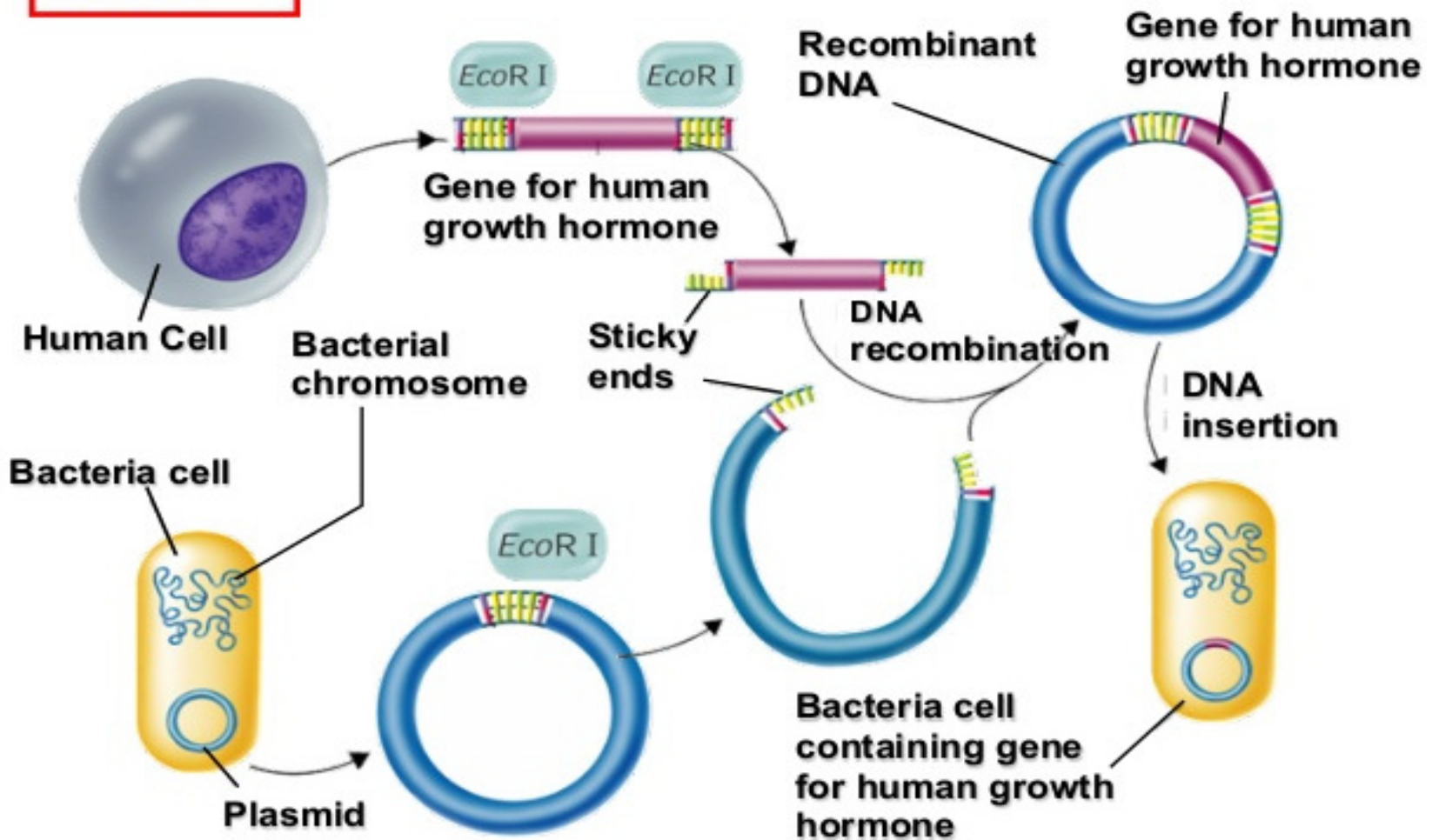
**E** **CO** **R** **I**

Gender Species Strain Enzyme  
number

**Figure 2.4.** Standardized nomenclature system for restriction enzymes.

# Basic Principle of Recombinant DNA Technology

**Movie**



# DNA Cloning

- Insert target DNA into a replicating cell
- Permits Amplification.
- “**Cloning Vector**” = Target Gene + Vector
- For example,
- The process of introducing foreign DNA into a cell
  - **Transformation** for bacteria
  - **Transfection** for eukaryotes.
- As Cell / Bacteria multiplies = Copies of clone DNA.
- Cloned DNA released from its vector by cleavage (using the appropriate restriction endonuclease) and is isolated.

# Vector

- It is DNA to which target gene is joined, for cloning.
- **Essential properties**
  1. **Autonomous replication**
  2. **At least one specific nucleotide sequence recognized by a restriction endonuclease in it**
  3. **At least one gene that confers the ability to select for the vector**
    - **such as an antibiotic resistance gene.**

Dr Piyush Tailor

# Vector

## 1. Plasmid

- Use to transfer 6-10 kbp gene

## 2. Bacteriophage

- Use to transfer 10-20 kbp gene

## 3. Cosmid

- Use to transfer  $> 20$  kbp gene

## 4. Bacterial Arteficial Chromosome (BAC)

Yeast Arteficial Chromosome (YAC)

Bacteriophage P1 Based Vector

- Several hundred kbp gene



# PLASMID

- Double stranded circular DNA.
- Easily trans-infect = easily enter from one bacteria to another bacteria
- Self Replicating.
- Antibiotic Resistant Property

Dr Piyush Tailor

# Bacteriophage & Cosmid

- Bacteriophage = Virus that infect bacteria
- Cosmids = Artificial constructs
- More efficiently
- Accommodate large DNA segments

Dr Piyush Tailor

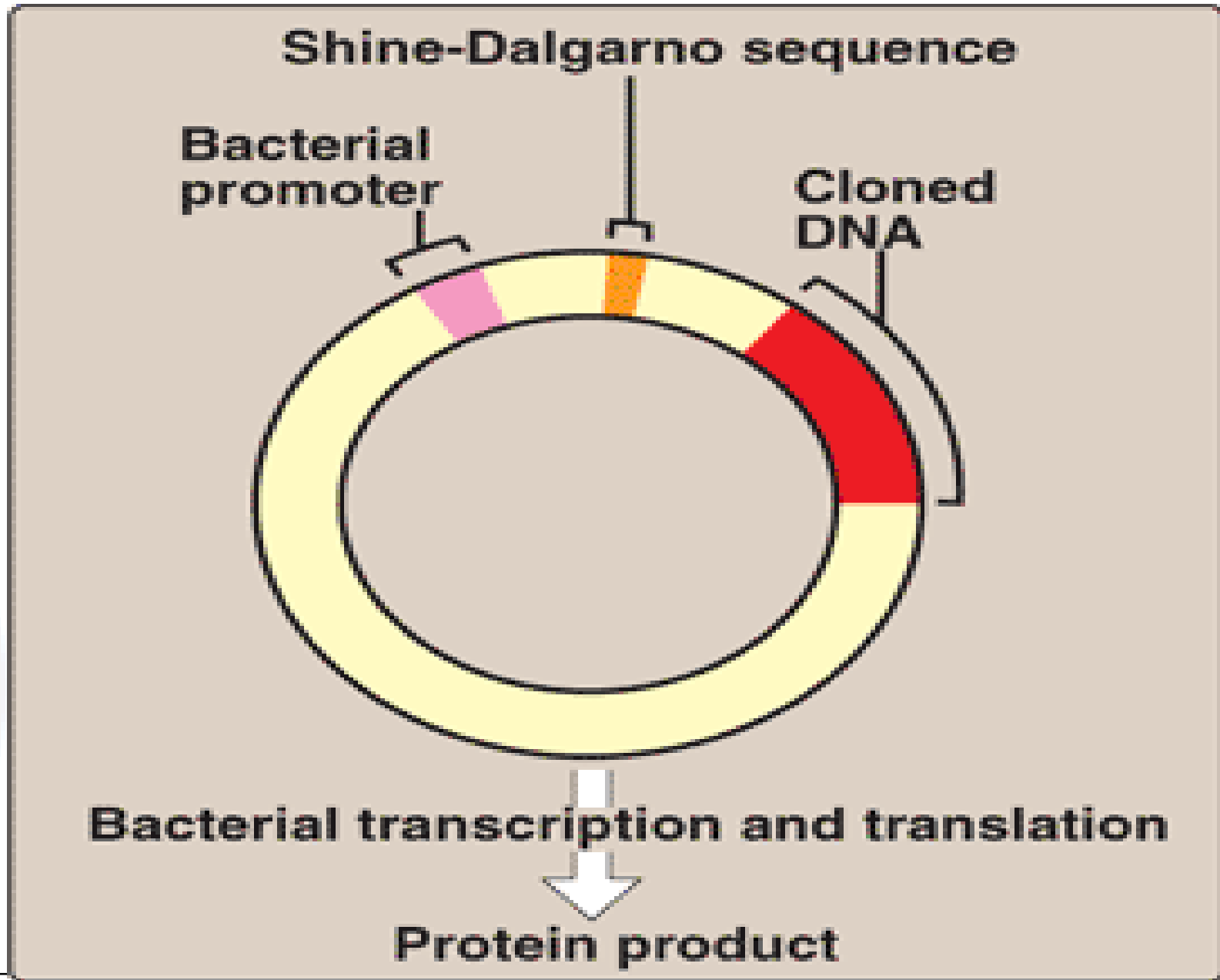
Is it only target gene require to clone,  
when we like to produce target protein  
(e.g. deficient enzyme)

from recombinant DNA technic

or

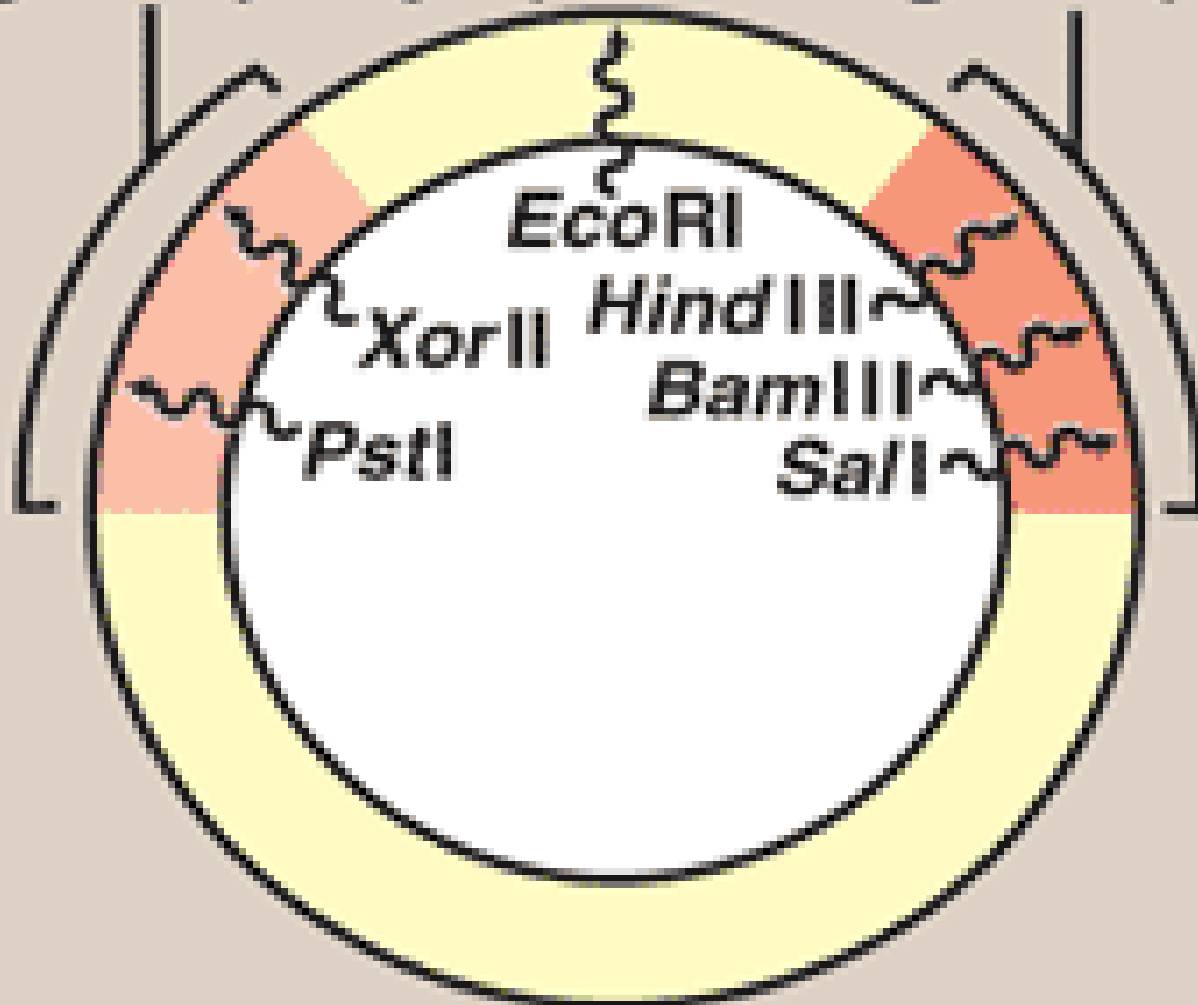
gene therapy?

# Expression Vector



Ampicillin  
resistance  
gene (*Amp<sup>R</sup>*)

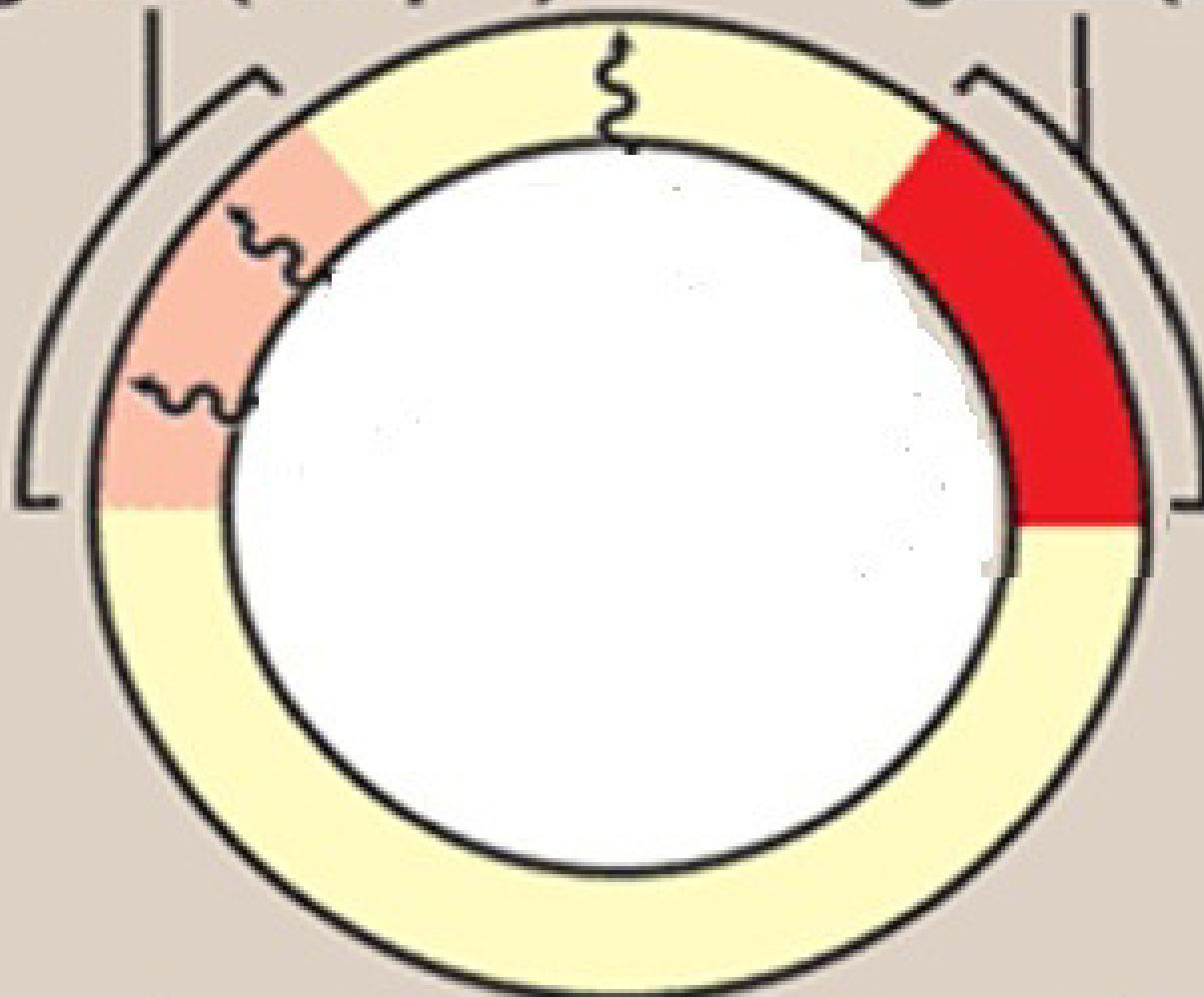
Tetracycline  
resistance  
gene (*Tet<sup>R</sup>*)



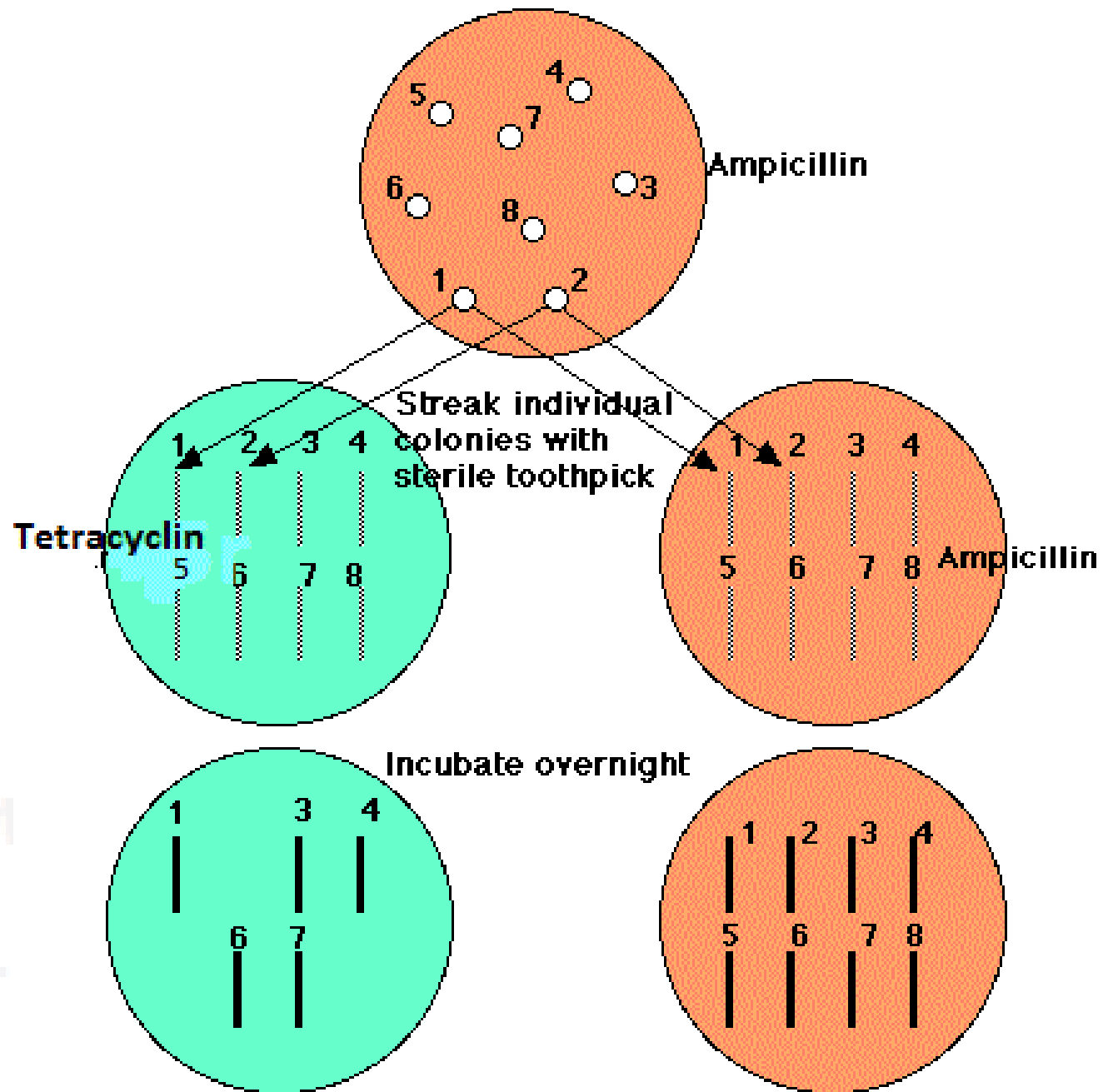
Plasmid pBR322

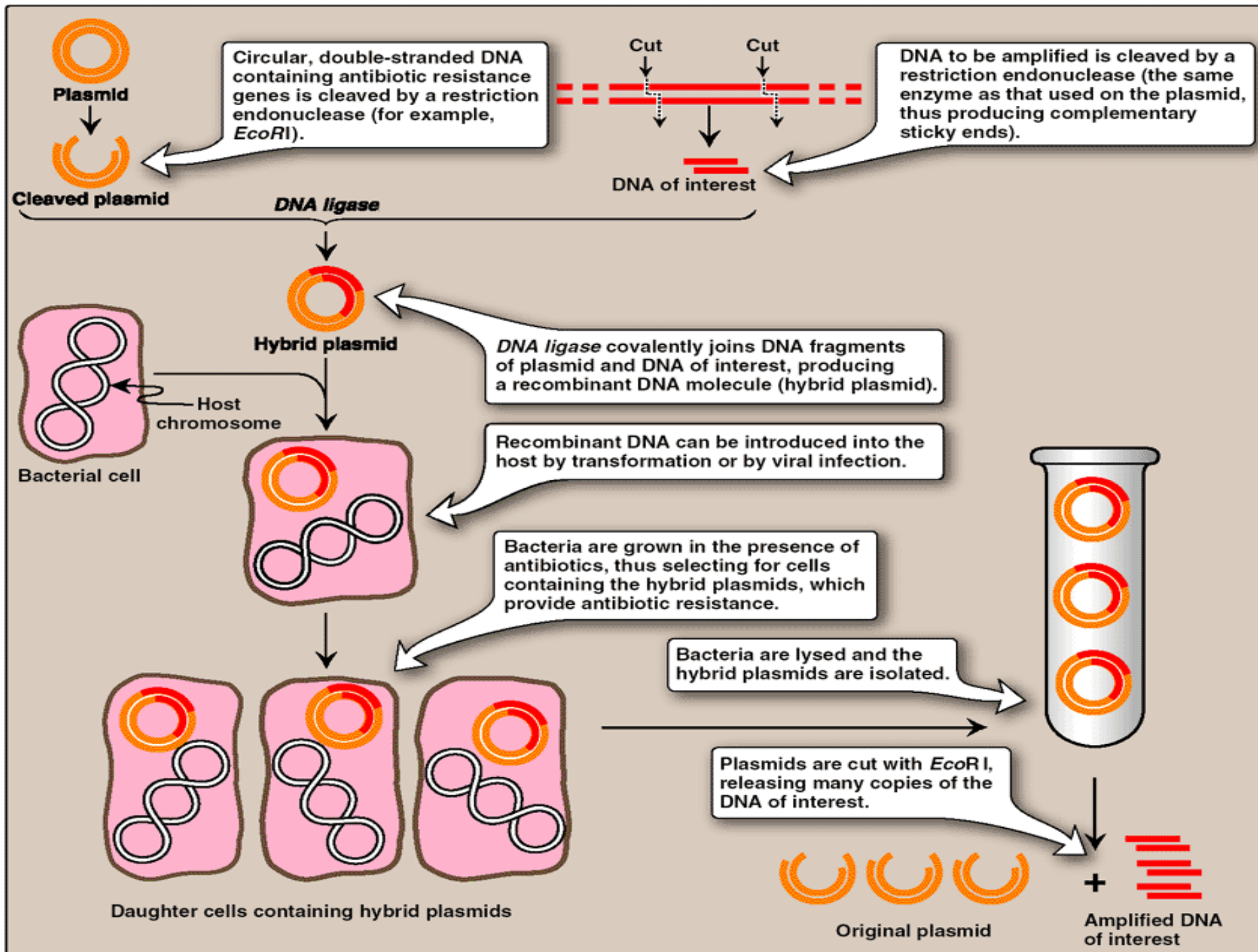
Ampicillin  
resistance  
gene (*Amp<sup>R</sup>*)

Target  
Gene



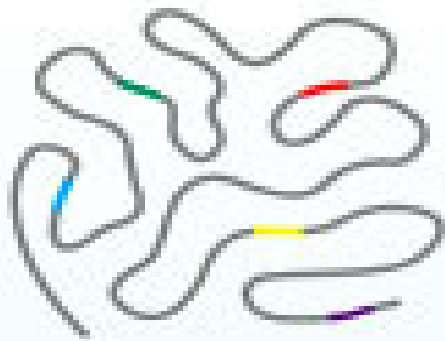
Chimeric DNA Plasmid







# Genomic Library



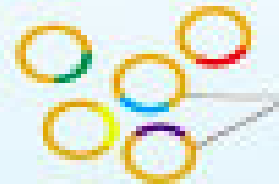
Human DNA

CLEAVE WITH  
RESTRICTION  
NUCLEASE



Millions of genomic  
DNA fragments

DNA FRAGMENTS  
INSERTED INTO  
PLASMIDS USING LIGASE



Recombinant  
DNA molecules

INTRODUCTION OF  
PLASMIDS INTO BACTERIA



Genomic library

# DNA Libraries

- Collection of cloned restriction fragments of the DNA

- Two type of libraries

## 1. Genomic libraries

- ✓ Copy of every DNA nucleotide sequence.
- ✓ Intron & Regulatory gene present.

## 2. Complementary DNA (cDNA) libraries.

- ✓ cDNA made from mRNA.
- ✓ Appear like mRNA molecules
- ✓ Intron & Regulatory genes absent.

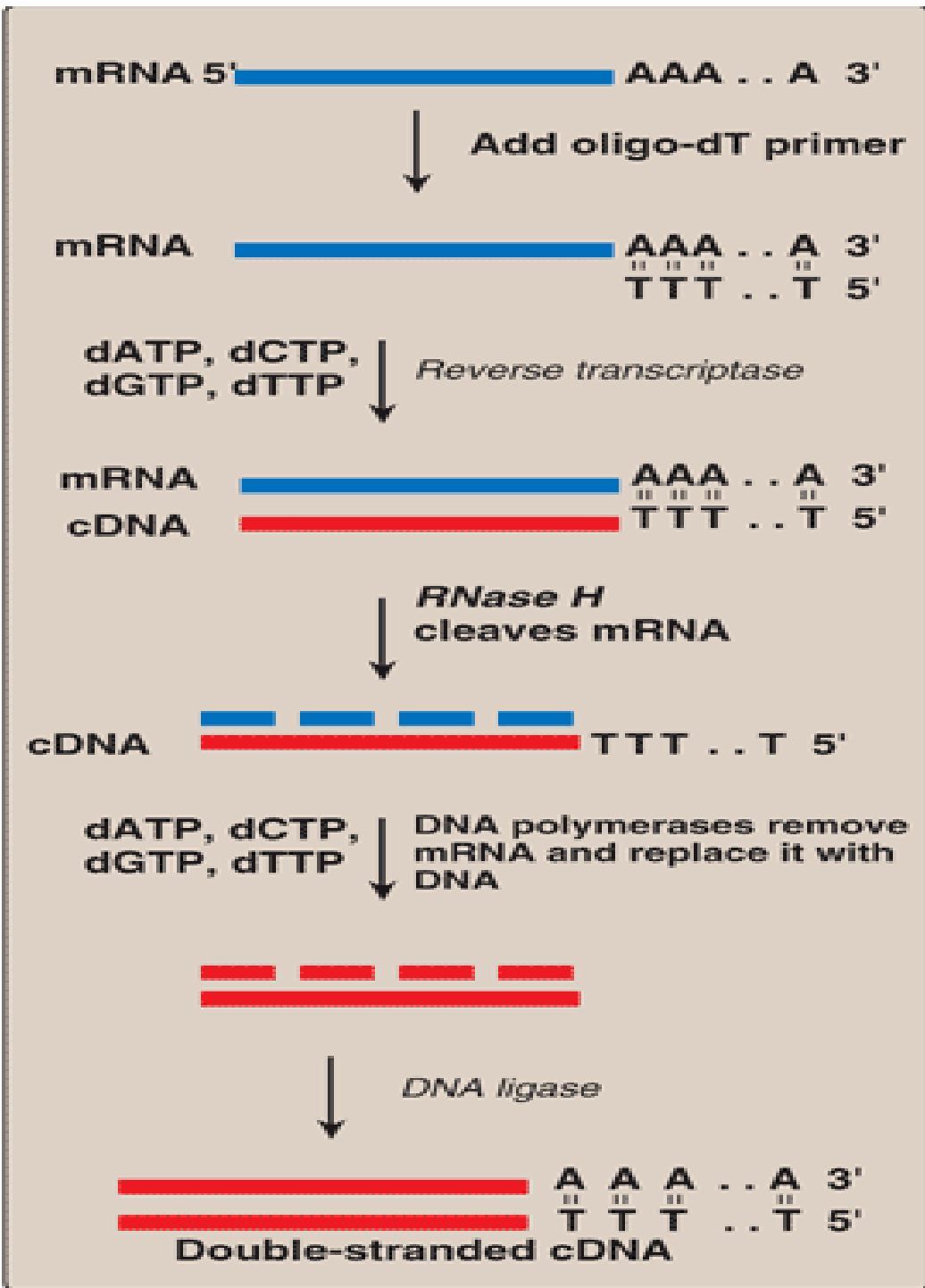
# Genomic DNA libraries

## Collection of fragments of total dsDNA

- Digestion of total DNA with restriction endonuclease
- Accordingly RE enzymes = more than one restriction site recognized
- If completion fragmentation
- Gene of interest is also fragmented
- No Gene for library.
- So, Partial digestion of DNA is performed
  - Limit - Amount of enzyme
  - Limit - Time of action of the enzyme.

# cDNA libraries

- Some gene of interest is expressed as mRNA
- Present at high concentrations in the cell.
- For example,
- Reticulocyte mRNA =  $\alpha$ -globin and  $\beta$ -globin.
- mRNA used as a template to dsDNA (cDNA).
- Using enzyme reverse transcriptase
- mRNA is isolated from tRNA & rRNA by the presence of its polyA tail.
- cDNA = amplified by cloning or PCR.
- Used as a probe to locate the gene that coded for the original mRNA
- cDNA has no intervening sequences



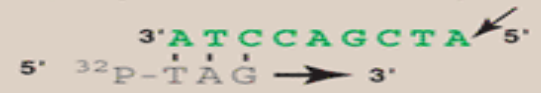
# Synthesis of cDNA from mRNA using Reverse Transcriptase

Tailor

# Sequencing of cloned DNA fragments

- *Sanger Dideoxy Method*
- ssDNA
- DNA polymerase.
- Radioactive primer complementary to the 3'-end of the target DNA
- Four deoxyribonucleoside triphosphates (dNTP).
- *Dideoxyribonucleoside triphosphates (ddNTP)* .
- Separation of DNA = Polyacrylamide gel electrophoresis
- Followed by autoradiography

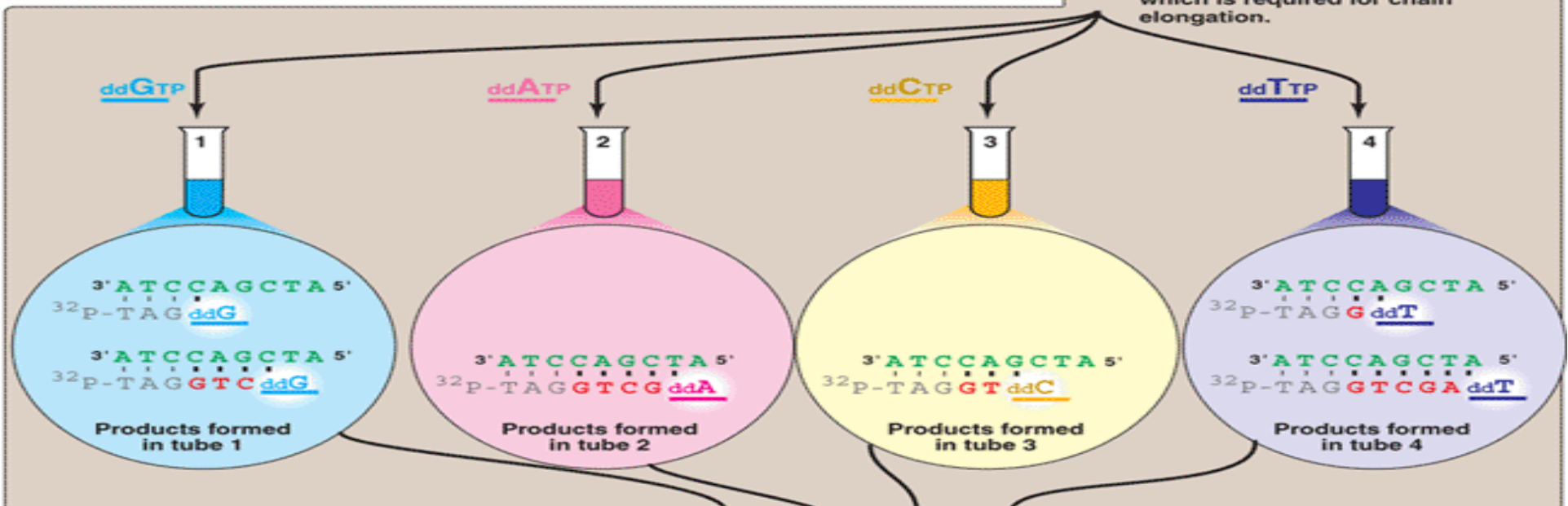
**1** Single-stranded DNA of unknown sequence is used as a template.



**2** Add primer, DNA polymerase + dATP, dGTP, dCTP, dTTP.

**3** Split the sample into four tubes, each with one of the four dideoxynucleotides.

**4** Synthesis proceeds until the dideoxynucleotide is incorporated into a DNA strand. DNA terminating in a dideoxynucleotide cannot be elongated because it lacks a 3'-OH, which is required for chain elongation.

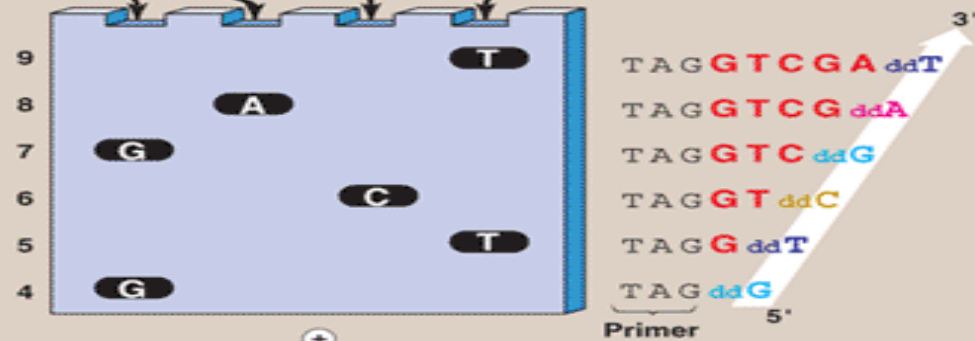


**5** Perform gel electrophoresis.

**6** Read the sequence of the newly synthesized strand (complementary to the original DNA sample).



Length of fragments (base pairs)



# Probes

## Use

- To find target DNA sequence from DNA fragments
- Used to identify which clone of library or which band on a gel contains target DNA.

## Characteristic

- Short sequence
- Single-stranded piece of DNA
- Labeled with a radioisotope or biotin.
- Complementary to the DNA of interest



# Hybridization of a probe to DNA fragments

- denatured of dsDNA = ssDNA
- Bound to solid support = nitrocellulose membrane.
- Hybridization by exogenous, single-stranded, radiolabeled DNA probe (complementary nucleotide sequence)
- Probes hinges to target DNA
- Extent of hybridization is measured by radioactivity.
- Excess Probe = Do not hybridize = Removed by washing.

Dr Piyush Tailor

# Synthetic oligonucleotide probes

- 20–30 nucleotides.
- Used to detect single-base changes

## cDNA probes

- Thousands of bases
- It can binding to a target DNA even with single-base change.

Dr Piyush Tailor

- Amino acid sequence of the protein may be used to construct a probe.
- Because of the degeneracy of the genetic code, it is necessary to synthesize several oligonucleotides.

Dr Piyush Tailor

# Allele-specific oligonucleotide (ASO) probe

**A**

**DNA from a patient with sickle cell disease**

**Portion of the gene for the  $\beta^S$ -chain of hemoglobin S**

DNA codes for valine instead of glutamate in the sixth position of  $\beta$ -globin.



Oligonucleotide probe hybridizes with a DNA fragment from the gene for the  $\beta$  chain of Hb S.

Probe

\*CTCCTGTTGGAGAAAGT

```
.. . . . GAGGACAACCTCTTTCAGACG .. . . .
```

DNA fragment coding for Hb S

## **B** DNA from a normal individual

Oligonucleotide probe fails to hybridize with the DNA fragment from the gene for the  $\beta$  chain of Hb A.

Probe

\*CTCCTG**T**GGAGAAGT

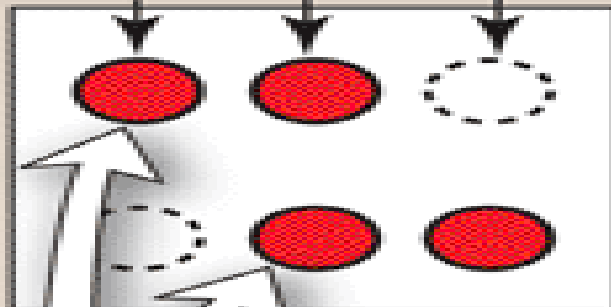
· · · GAGGACT**T**CCTCTTCAGACG · · ·

DNA fragment  
coding for Hb A

Homozygous normal, (AA)

Heterozygous (carrier, AS)

Homozygous mutant (affected, SS)



This row was probed with an ASO specific for a normal  $\beta^A$  gene

This row was probed with an ASO specific for a mutant  $\beta^S$  gene

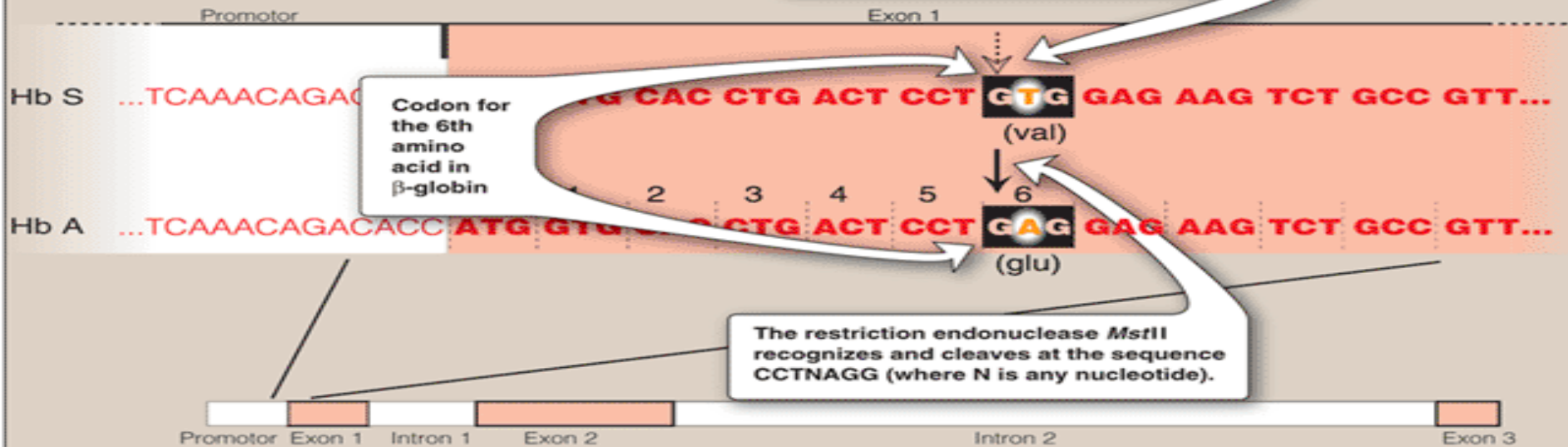
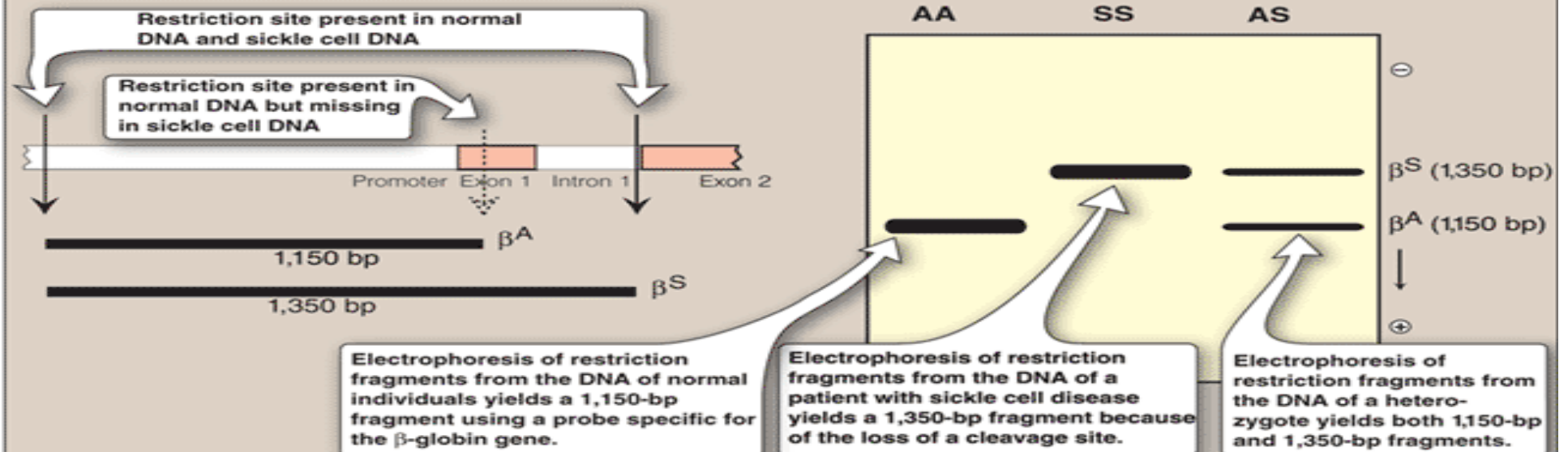


= Probe hybridizes with patient's DNA

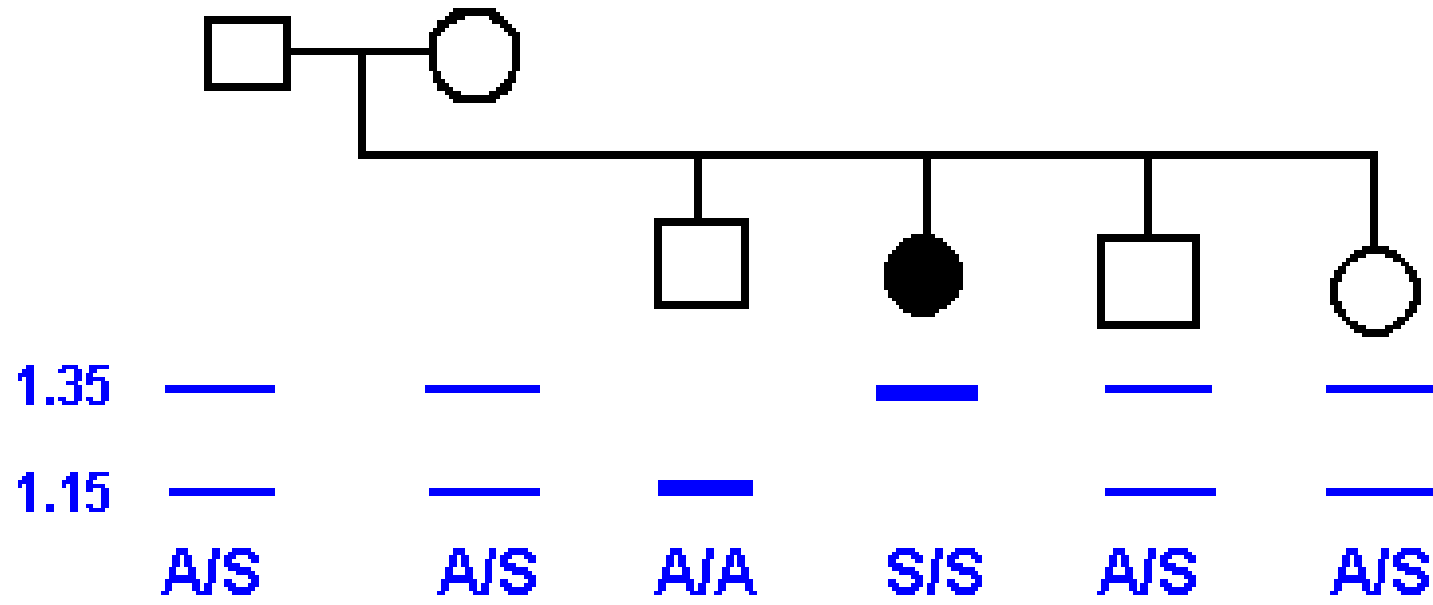
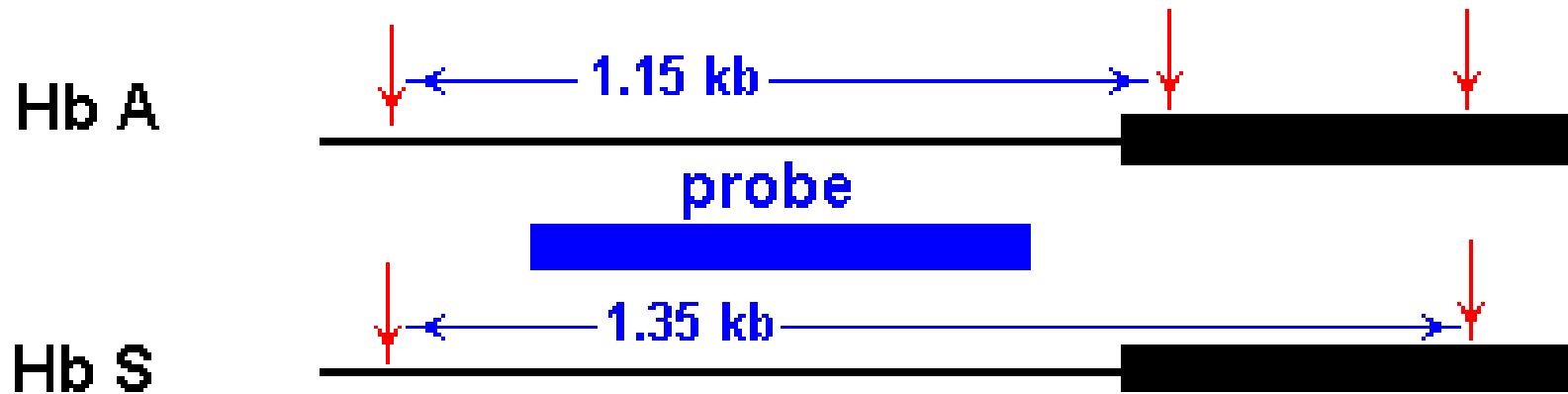


= Probe does not hybridize to patient's DNA

Two samples of DNA from each individual are applied to the membrane.

**A****Details of a portion of the  $\beta$ -globin gene****B****Cleavage of the  $\beta$ -globin gene with a restriction endonuclease**

### restriction site locations





# Biotinylated probes

- **Disposal of radioactive = expensive.**
- **Biotin + 4 Avidin ( Egg-Protein )**
- Avidin can be attached to a fluorescent dye
- Great sensitivity.

Dr Piyush Tailor

# Antibodies

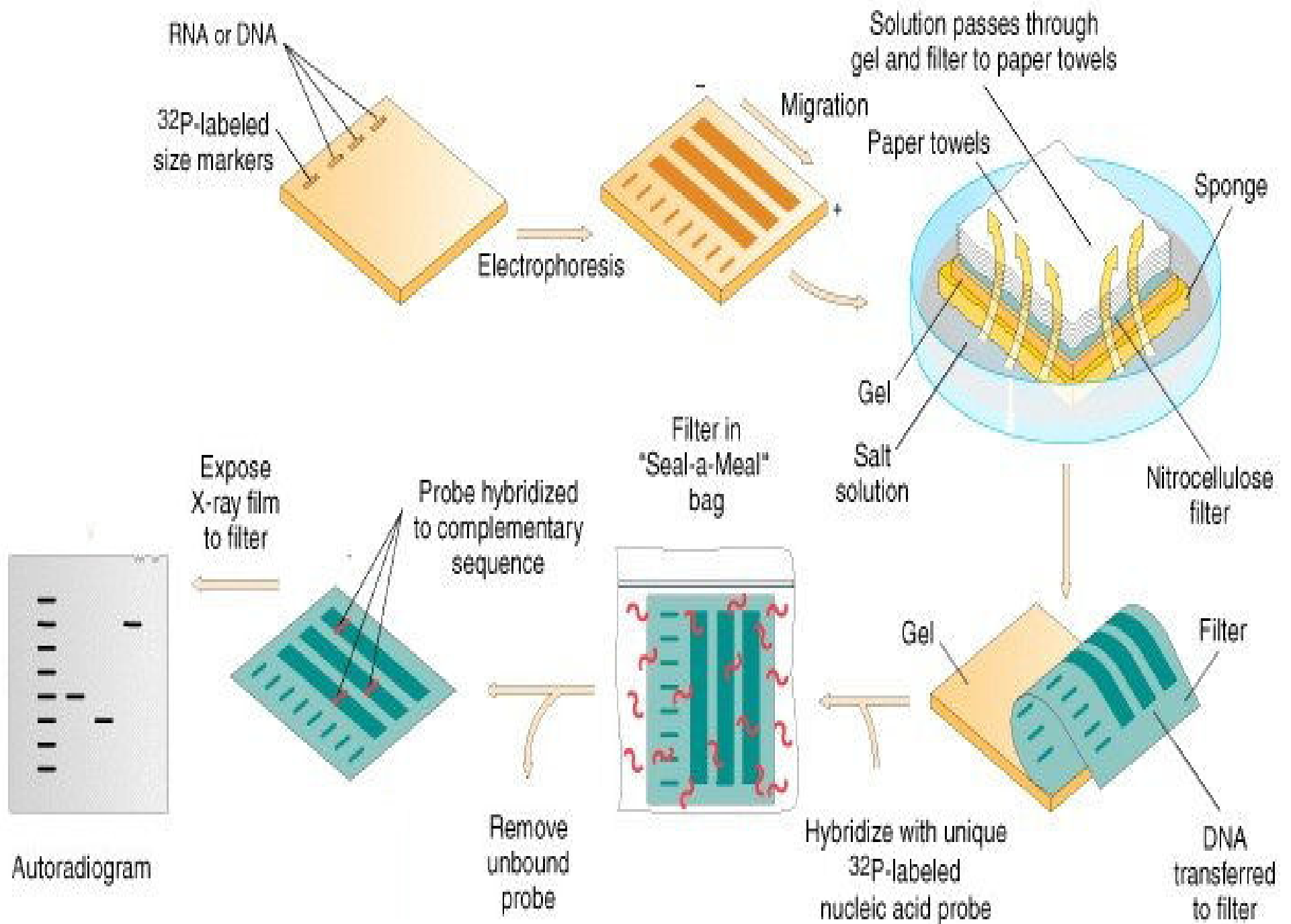
- When amino acid sequence is unknown for synthesis of a probe.
- A labeled antibody is used to identify which bacterial colony produces the protein
- Gene can be identified indirectly by cloning cDNA in an expression vector that allows the cloned cDNA to be transcribed and translated.
- [Note: A library created using expression vectors is called an expression library.]

Dr. Piyush Tailor

# (Edward)Southern Blotting

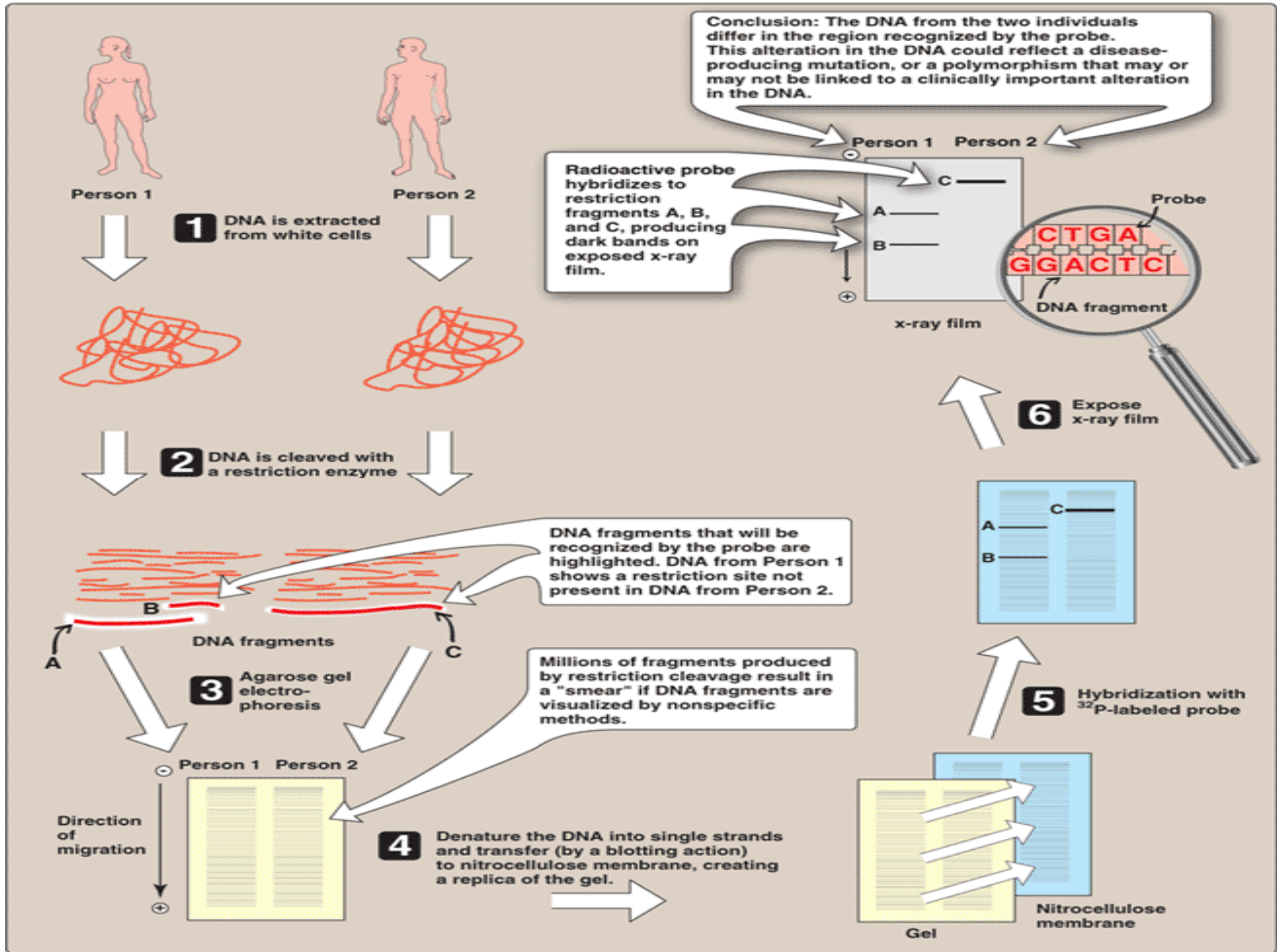
- “Southern” = Detect in DNA.
- “Northern” = mRNA
- “Western” = Protein
- It combines the use
  - Restriction enzymes
  - Electrophoresis
  - DNA probes
  - Blotting – Nitrocellulose paper

Dr Piyush Tailor



# Southern Blot Procedure

- DNA is **extracted** from cells
- **Cleaved** into many fragments = **Restriction Enzyme**.
- Separated on basis of size by **Electrophoresis**.
- **Denatured and transferred (blotted)** to a nitrocellulose membrane
- If Whole DNA = Millions of Fragment copy
- Blure & Overlapping bands.
- To avoid this = Uses a **probe** to identify Target DNA.
- Expose to **X – ray** film
- **Comparison** of the position of the band to standard fragments.
- Band pattern depend = on Restriction Endonuclease  
= on Probe



# Detection of mutations

- Mutation = Pattern of bands is different.
- *Sometime Mutation may not affect a restriction site, with one specific restriction enzyme.*
- *It may be done by using a different restriction enzyme, those can recognize sequence affected by mutation.*

Dr Piyush Tailor

# Restriction Fragment Length Polymorphism

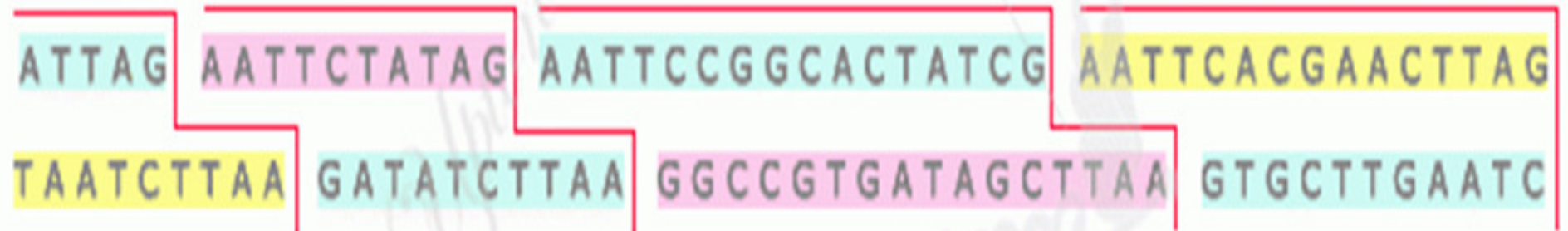
- Genomic variation in DNA among more than 1% of population.
- Differ in 0.1% of genome.
- Genome variations include both polymorphisms and mutations.
- Polymorphism = not always harmful
- Mutation = harmful
- Polymorphisms = in the Introns = that do not code for proteins.
- RFLP can be examined by cleaving DNA into fragments with R.E.
- Length of restriction fragments is altered if the genetic variant alters the DNA

Dr Piyush Tailor





Using EcoRI to cut example DNA at GAATTC produces the following four strand fragments.

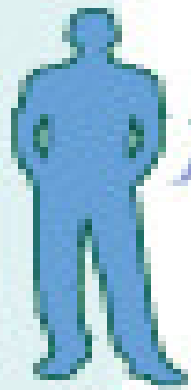


# DNA variations resulting in RFLP

- Two types of DNA RFLP
  1. Single-base changes
  2. Tandem Repeats

Dr Piyush Tailor

# SNP



G C A **A** C G T T A G A



G C A **G** C G T T A G A



G C A **T** C G T T A G A

# Single Nucleotide Polymorphism (SNP)

Individual 1

Chr 2 . . . CGATATTCCTATCGAATGTC . . .  
*copy1* . . . GCTATAAGGATAGCTTACAG . . .

Chr 2 . . . CGATATTC~~C~~ATCGAATGTC . . .  
*copy2* . . . GCTATAAGGGTAGCTTACAG . . .

Individual 2

Chr 2 . . . CGATATTC~~C~~ATCGAATGTC . . .  
*copy1* . . . GCTATAAGGGTAGCTTACAG . . .

Chr 2 . . . CGATATTC~~C~~ATCGAATGTC . . .  
*copy2* . . . GCTATAAGGGTAGCTTACAG . . .

Individual 3

Chr 2 . . . CGATATTCCTATCGAATGTC . . .  
*copy1* . . . GCTATAAGGATAGCTTACAG . . .

Chr 2 . . . CGATATTCCTATCGAATGTC . . .  
*copy2* . . . GCTATAAGGATAGCTTACAG . . .

Individual 4

Chr 2 . . . CGATATTCCTATCGAATGTC . . .  
*copy1* . . . GCTATAAGGATAGCTTACAG . . .

Chr 2 . . . CGATATTC~~C~~ATCGAATGTC . . .  
*copy2* . . . GCTATAAGGGTAGCTTACAG . . .

Individual 5

Chr 2 . . . CGATATTC~~C~~ATCGAATGTC . . .  
*copy1* . . . GCTATAAGGGTAGCTTACAG . . .

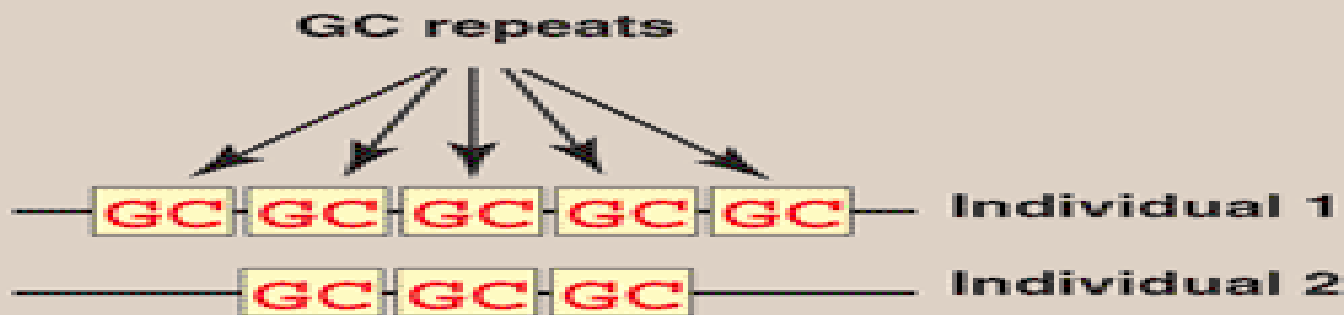
Chr 2 . . . CGATATTCCTATCGAATGTC . . .  
*copy2* . . . GCTATAAGGATAGCTTACAG . . .

Individual 6

Chr 2 . . . CGATATTC~~C~~ATCGAATGTC . . .  
*copy1* . . . GCTATAAGGGTAGCTTACAG . . .

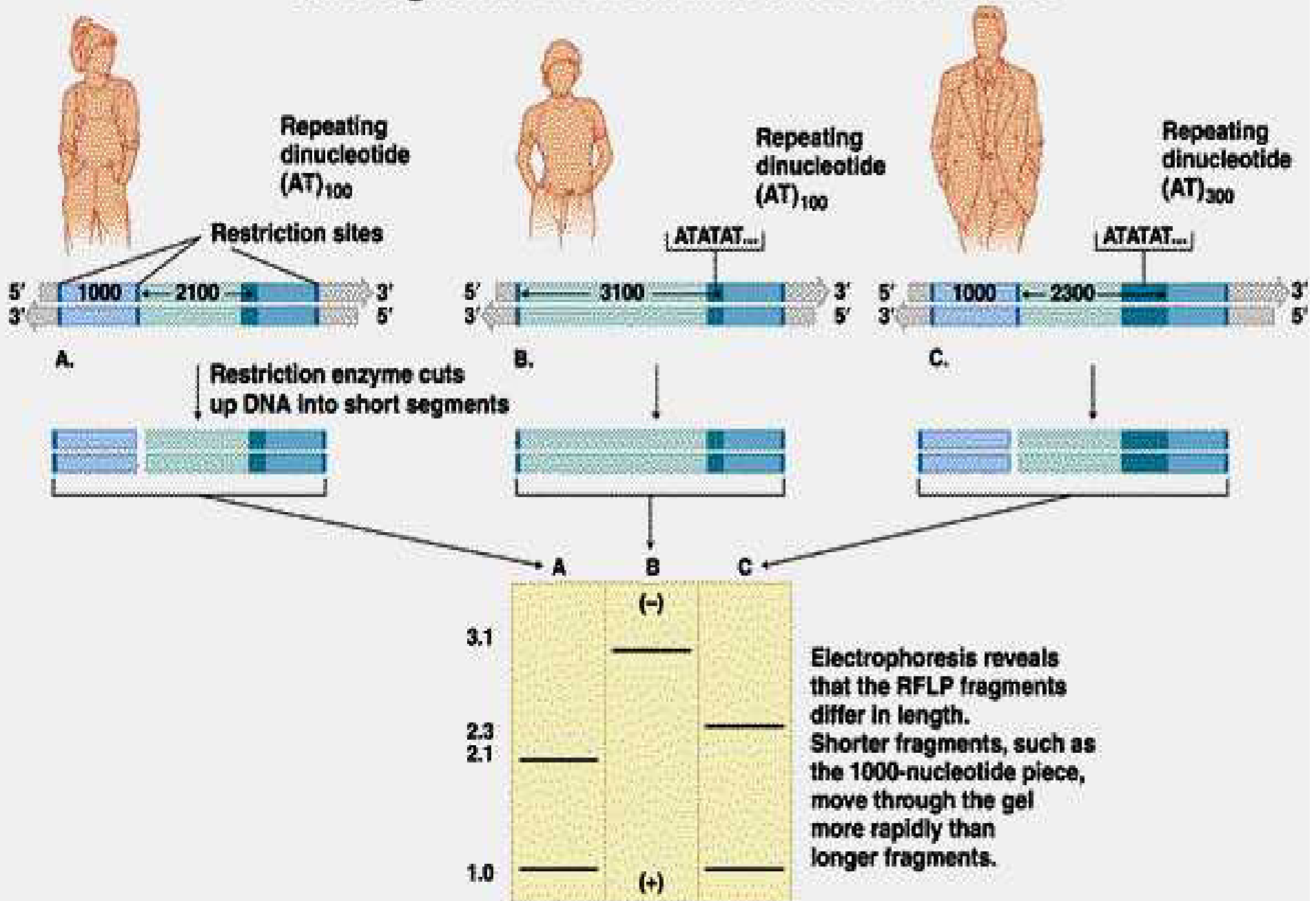
Chr 2 . . . CGATATTCCTATCGAATGTC . . .  
*copy2* . . . GCTATAAGGATAGCTTACAG . . .

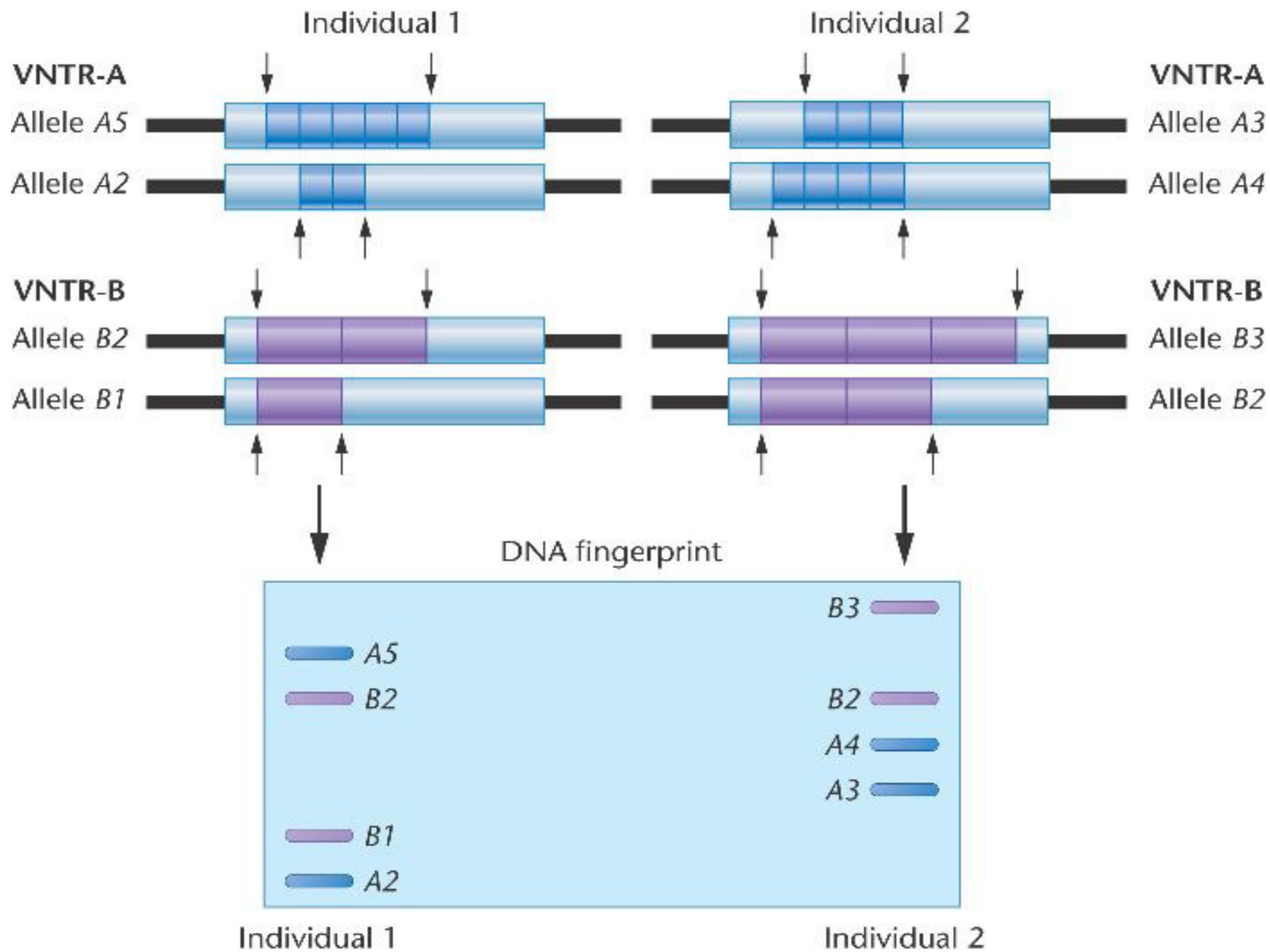
Polymorphisms can occur either in the sequence of bases at a single nucleotide locus (called SNP if only one base is altered) or ...

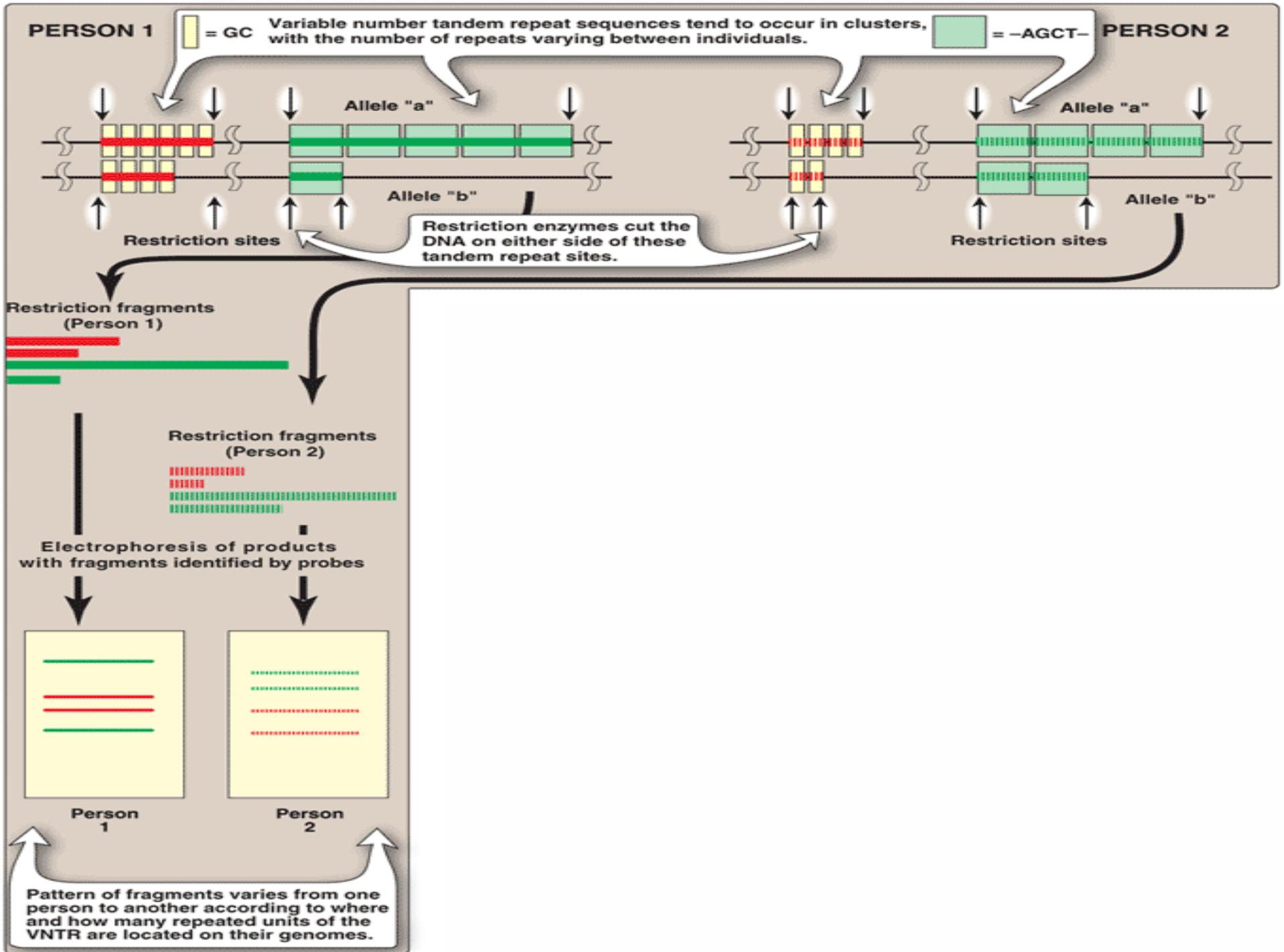


... there can be polymorphisms where variable numbers of tandem repeats (VNTR) occur. A specific number of tandem repeats defines a VNTR allele at a particular locus.

# Same segment of DNA from three different individuals:









# Single base changes in DNA

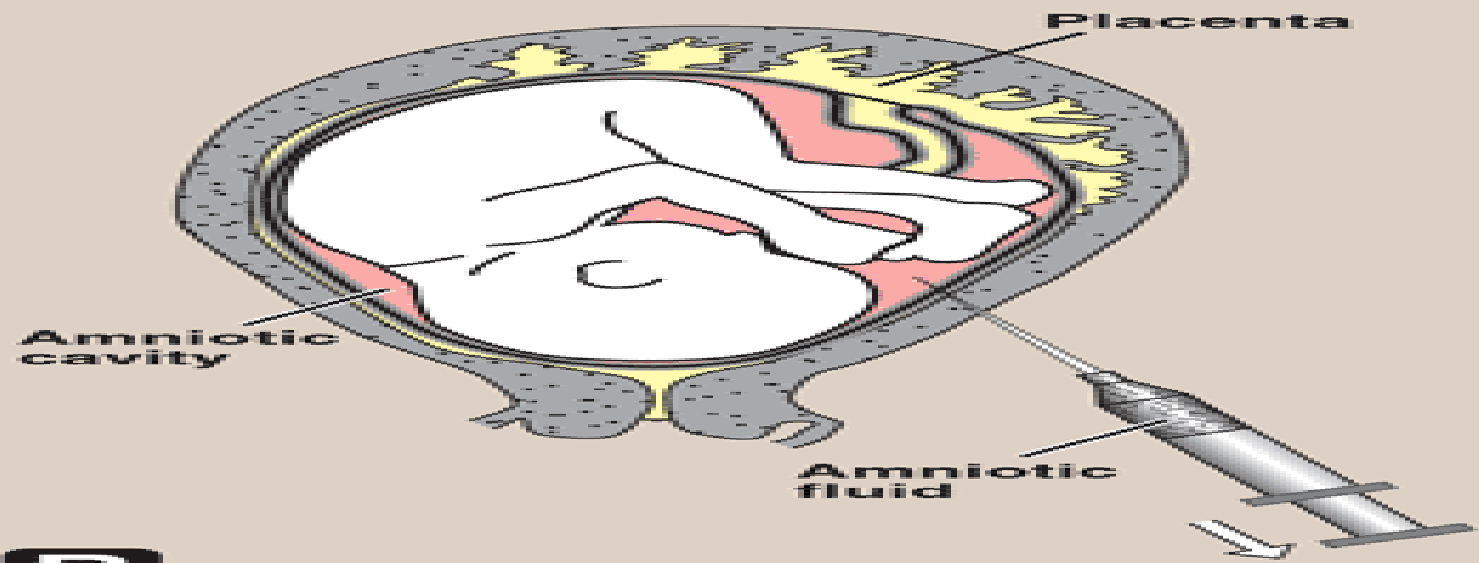
- About 90%
- SNPs = Single Base Polymorphism (“snips”)
- New restriction site created
- Results in fragments of lengths differing from the normal
- Detected by DNA hybridization

Dr Piyush Tailor

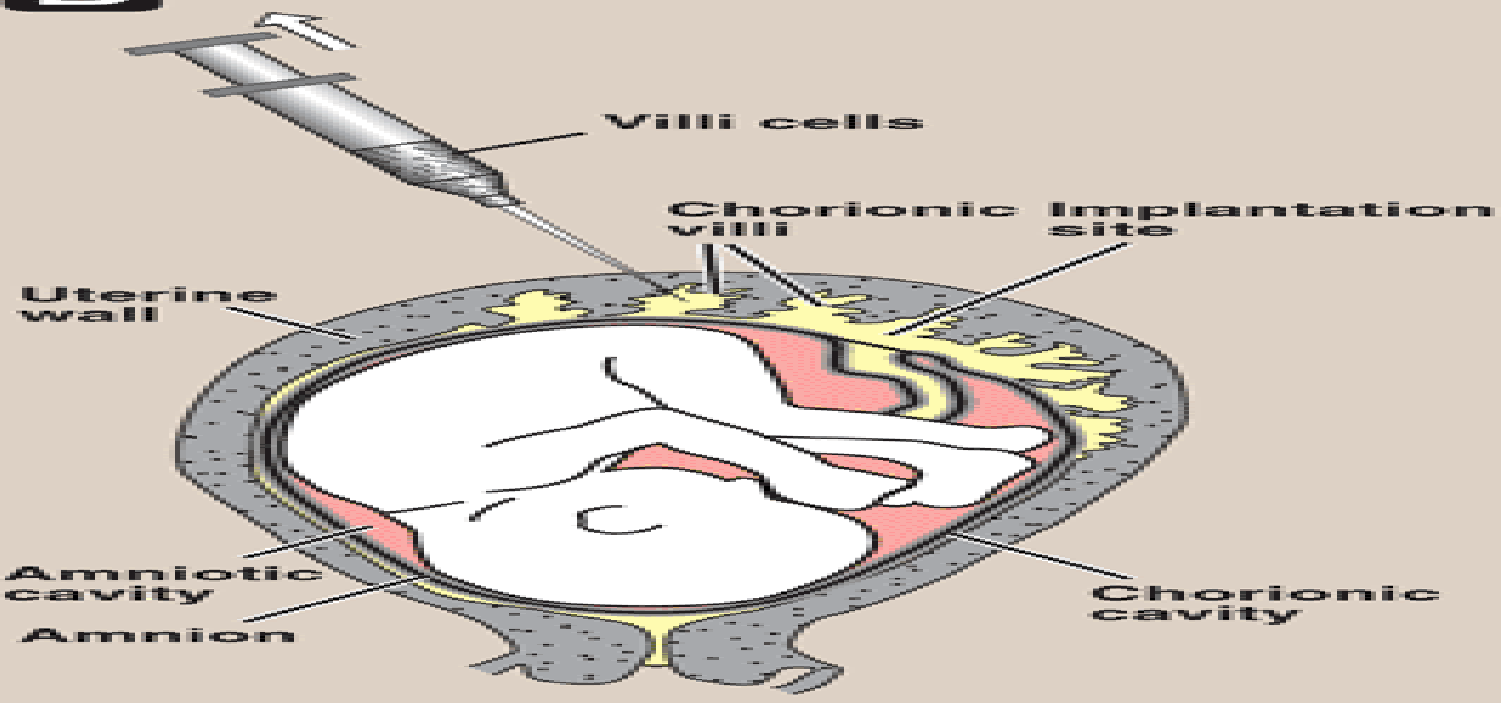
# Tandem Repeats

- Variable number of tandem repeats (VNTR)
  - Short sequences
  - Scattered location
  - Repeated in tandem (one after another).
  - Number of VNTR units varies from person to person
  - unique for any given individual.
- Useful for DNA fingerprint
- In forensic and paternity identity cases.
- No known effect on the structure or rate of production of any particular protein.

**A**

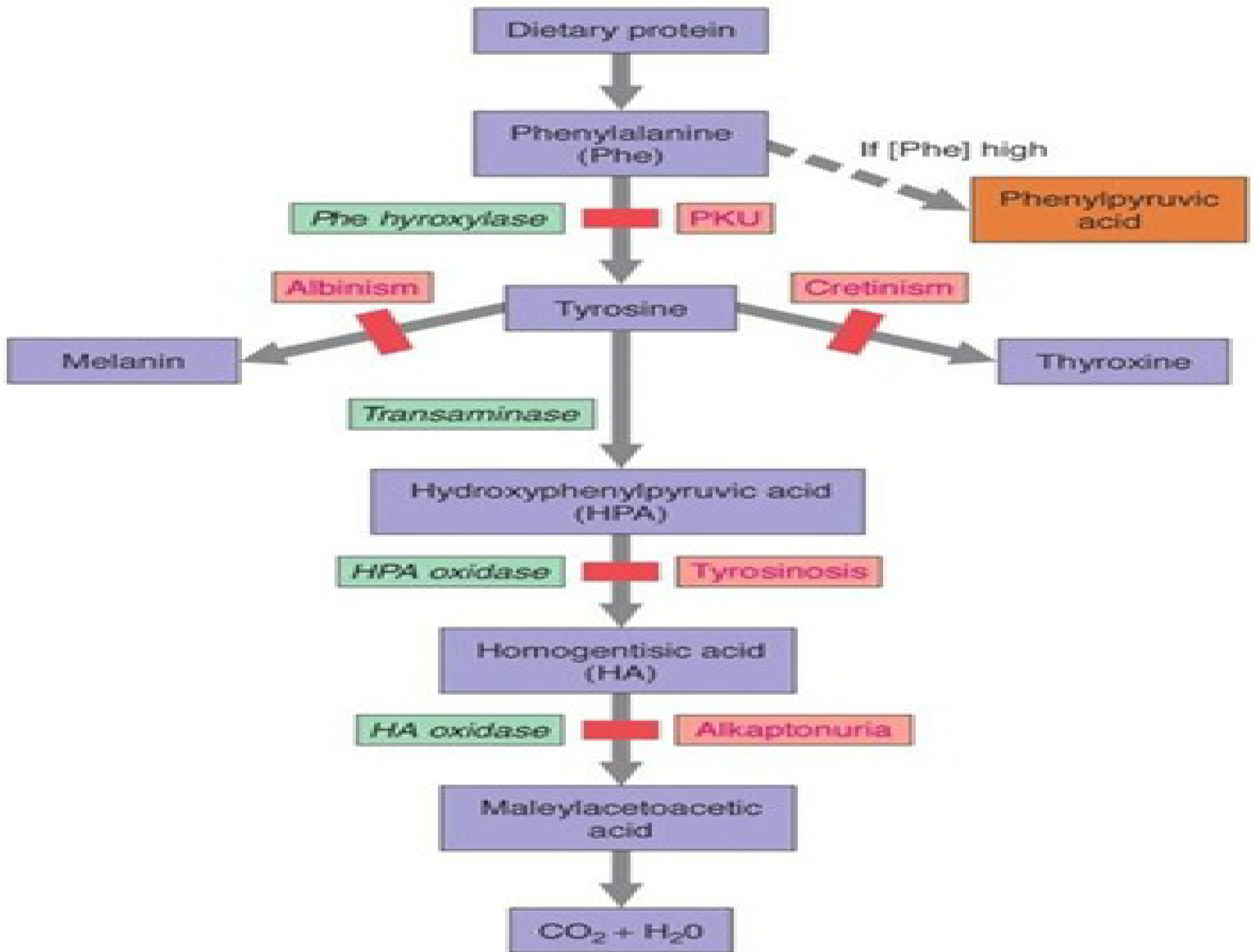


**B**



# Pre-natal diagnosis by RFLP

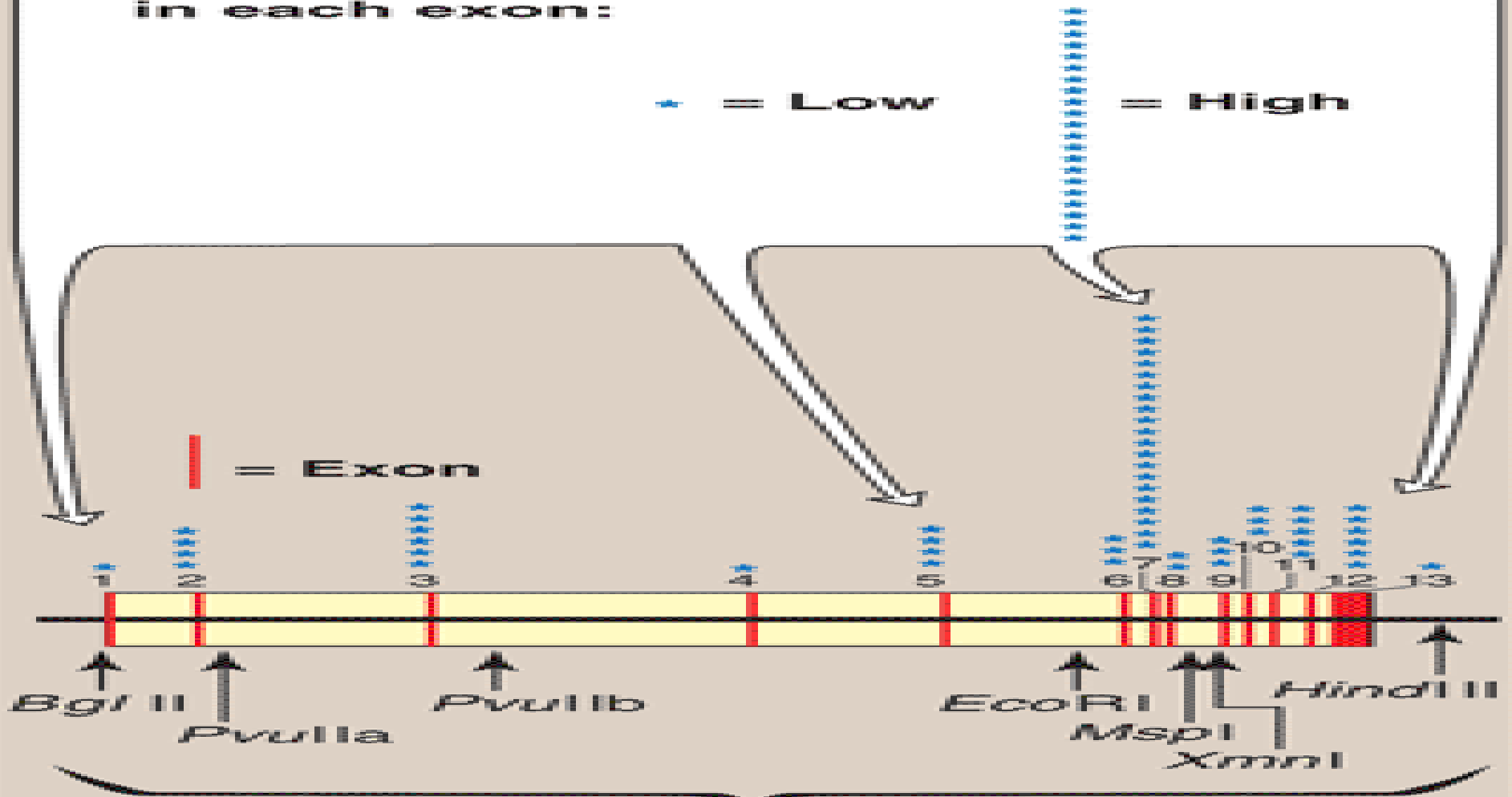
- **Direct diagnosis** : Sickle cell disease
- **Indirect diagnosis** : Phenylketonuria
  - Phenylalanine hydroxylase (PAH) gene deficient
  - On chromosome 12
  - 90kb of genomic DNA & 13 exons
  - Mutations in this gene usually do not directly affect any restriction endonuclease recognition site.
  - To establish a diagnostic protocol for this genetic disease, one has to analyze DNA of family members of the afflicted individual. The key is to identify markers (RFLP) that are tightly linked to the disease trait.
  - Once these markers are identified, RFLP analysis can be used to carry out prenatal diagnosis.



Mutations in the *phenylalanine hydroxylase* gene occur in all thirteen exons of the gene. The majority are missense mutations, although splice, nonsense, and silent mutations, as well as deletions and insertions, have been found.

Relative frequency of mutation in each exon:

★ = Low      = High



Some sites cleaved by restriction enzymes

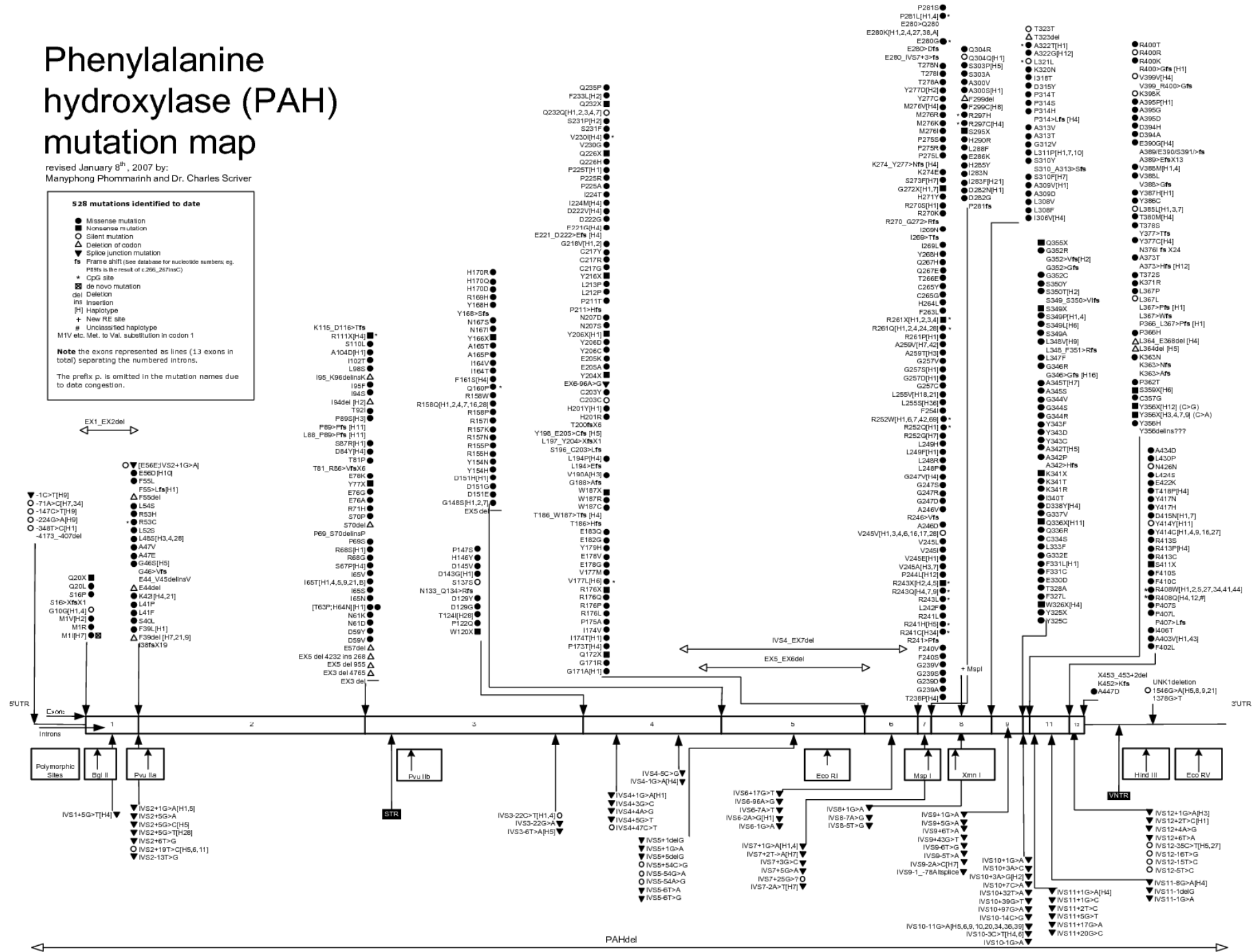
# Phenylalanine hydroxylase (PAH) mutation map

revised January 8<sup>th</sup>, 2007 by:  
Manyphong Phommahnh and Dr. Charles Scriver

## 528 mutations identified to date

- Missense mutation
- Nonsense mutation
- Silent mutation
- △ Deletion of codon
- ▽ Splice junction mutation
- fs Frame shift (See database for nucleotide numbers; eg. P99fs is the result of c.266\_268ins/c)
- ⊕ CpG site
- ⊕ de novo mutation
- del Deletion
- ins Insertion
- [H] Haplotype
- + New RE site
- # Unclassified haplotype
- M1V etc. Met. to Val. substitution in codon 1

Note the exons represented as lines (13 exons in total) separating the numbered introns.  
The prefix p. is omitted in the mutation names due to data congestion.



PAHdel

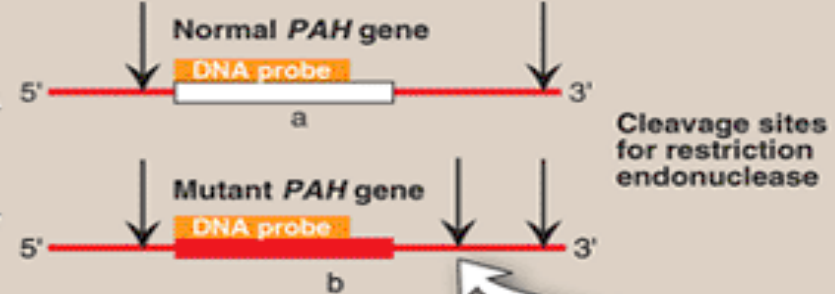
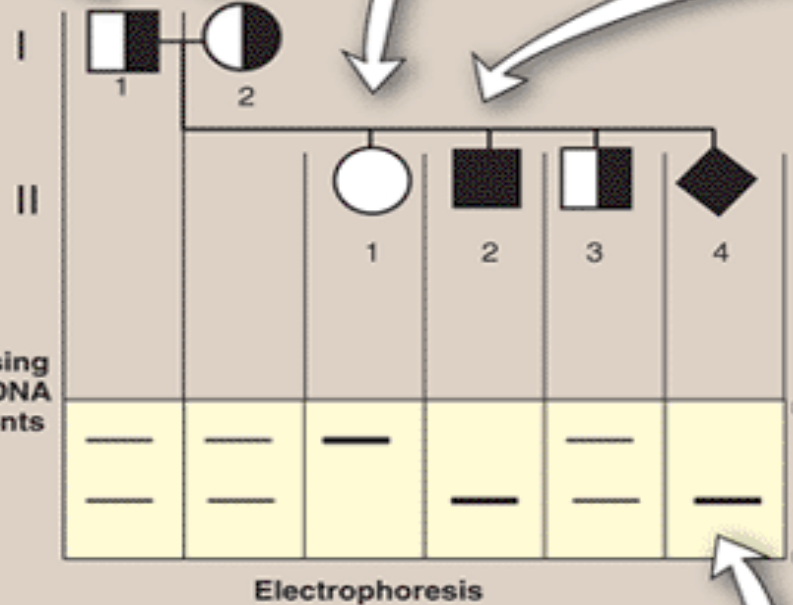
The parents are both heterozygous for the phenylalanine hydroxylase gene. They, thus, have both fragment "a" (normal) and fragment "b" (defective) cleaved by a restriction nuclease.

This child is unaffected and shows only fragment "a" when DNA is digested with the same restriction endonuclease as her parents' DNA. Thus, the normal gene is associated with the polymorphism giving fragment "a."

This child is affected (lacks *phenylalanine hydroxylase* activity) and shows only fragment "b" when DNA is digested with the same restriction endonuclease. Thus, the defective gene is associated with the polymorphism giving fragment "b."

A radioactive probe hybridizes with DNA fragments "a" and "b".

Decreasing size of DNA fragments  
↓



Fetal DNA shows only fragment "b" when DNA is digested with the same restriction endonuclease. This means that the fetus is affected because it has inherited two abnormal genes from its parents and shows the genotype "bb."

The presence of a polymorphic site permits cleavage by the enzyme and, therefore, yields fragment "b." [Note: The polymorphic site is, in general, not the structural alteration that causes the disease and in this case is not even in the coding part of the gene.]

Legend:

- Male
- Female
- Fetus
- Heterozygote
- Homozygote

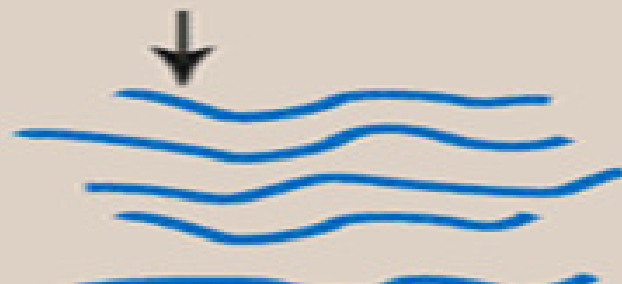
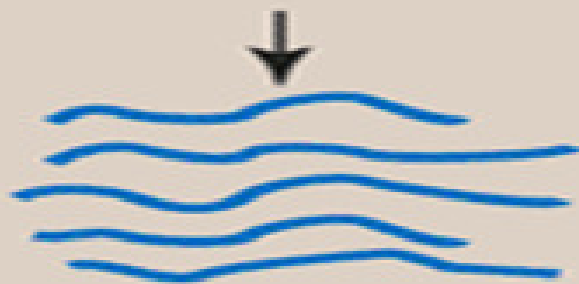
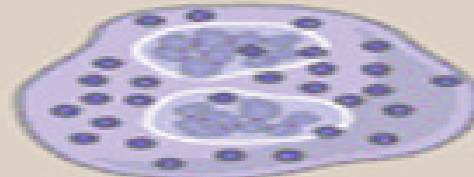
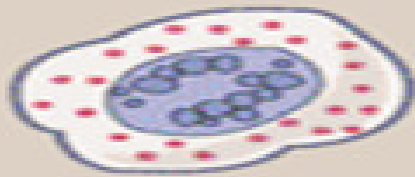


# Microarray

- Contain **thousands of immobilized DNA** sequences
- Used to analyze a sample for the gene variations & mutation
- **Analyzing thousands of genes** at the same time.
- mRNA = converted to cDNA
- labeled with a fluorescent tag
- This mixture is then exposed to a gene chip, which is a glass slide or membrane containing thousands of tiny spots of DNA, each corresponding to a different gene.
- Amount of fluorescence is measure

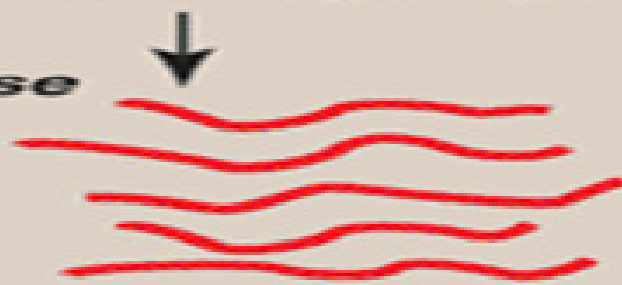
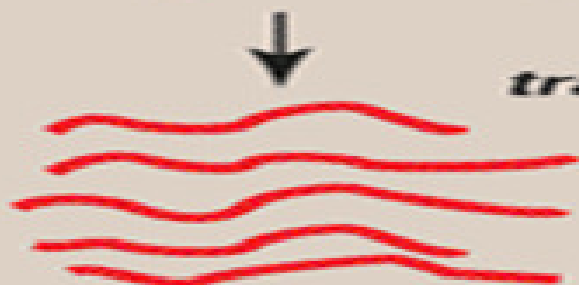
**Normal cell**

**Cancer cell**



**mRNA**

*Reverse transcriptase*

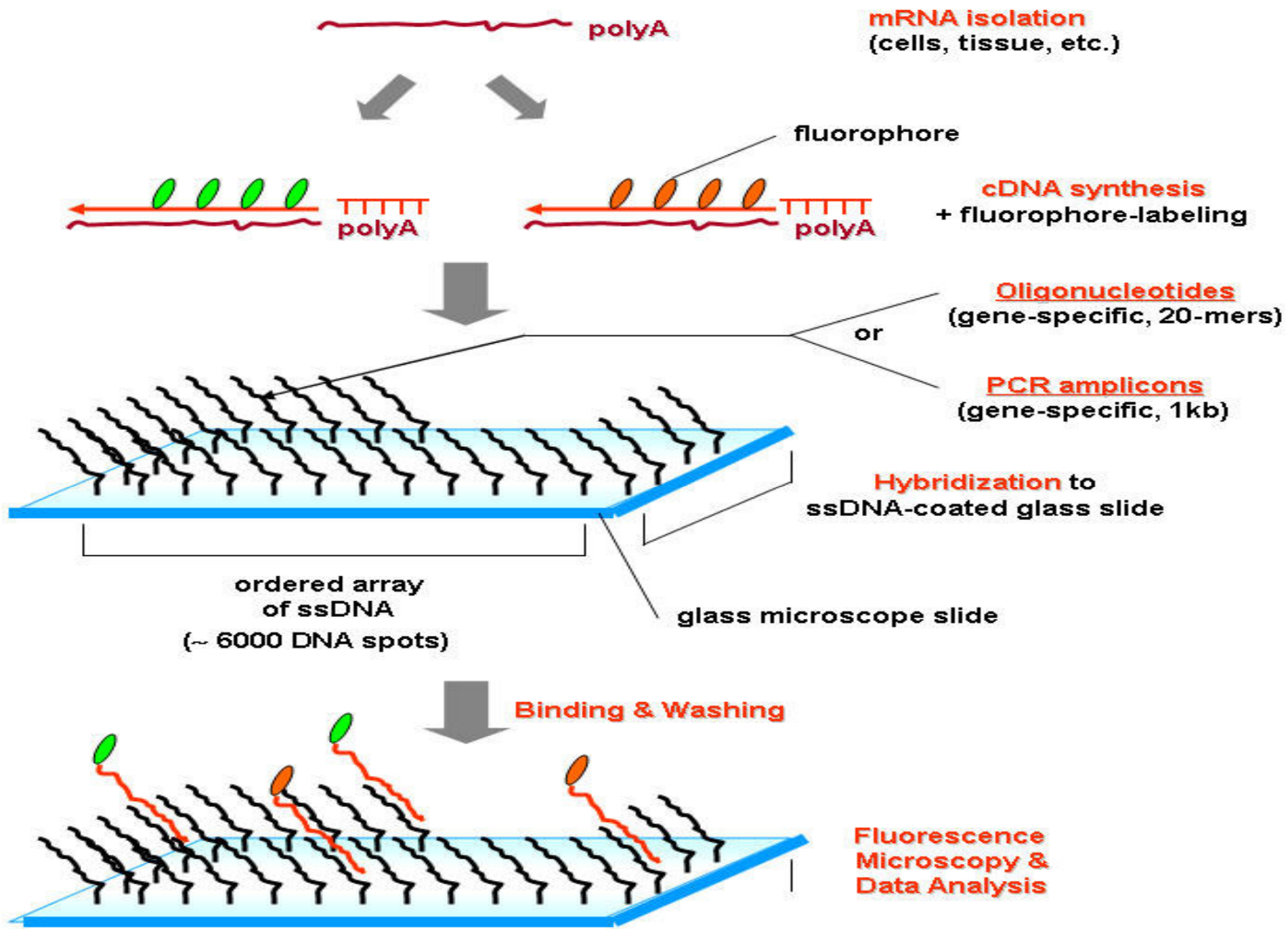


**cDNA**

**Label with green fluorescent molecule**

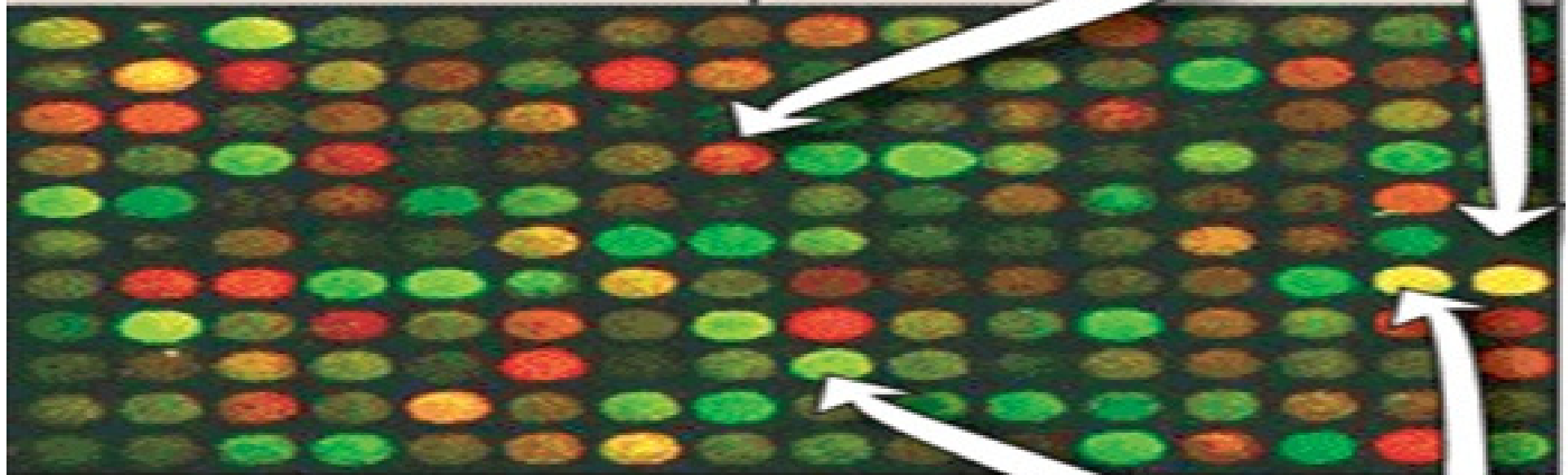
**Label with red fluorescent molecule**

**Mix cDNAs and hybridize to microarray**



**Dark (black) spot: Neither cell produces this message.**

**Red spot: Cancer cell produces more of its message.**



**Green spot: Normal cell produces more of its message.**

**Yellow spot: Both cells produce the same amount of message.**

**Patient 1**

**Patient 2**

**Patient 3**

**Test blood samples by ELISA assay**



**Positive**

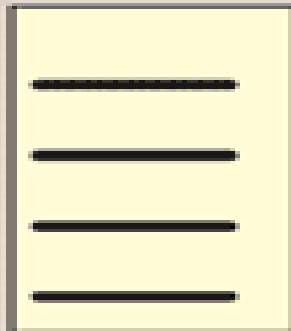


**Positive**



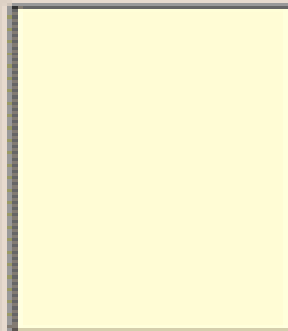
**Negative**

**Retest using Western blots**



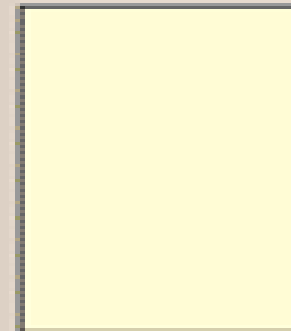
**Positive**

**The patient's serum contains antibodies to HIV**



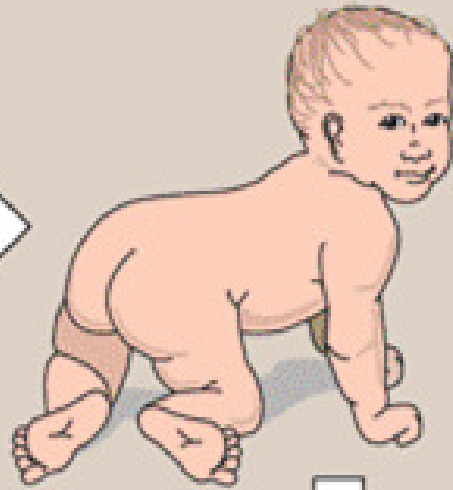
**Negative**

**The patient's serum gave a false-positive response in the ELISA assay**



**Negative**

<b>TECHNIQUE</b>	<b>SAMPLE ANALYZED</b>	<b>GEL USED</b>	<b>PURPOSE</b>
Southern blot	DNA	Yes	Detects DNA changes
Northern blot	RNA	Yes	Measures mRNA amounts and sizes
Western blot	Protein	Yes	Measures protein amounts
ASO	DNA	No	Detects DNA mutations
Microarray	RNA or cDNA	No	Measures many mRNA levels at once
ELISA	Proteins or antibodies	No	Detects proteins (antigens) or antibodies
Proteomics	Proteins	Yes	Measures abundance, distribution, posttranslational modifications, functions, and interactions of cellular proteins



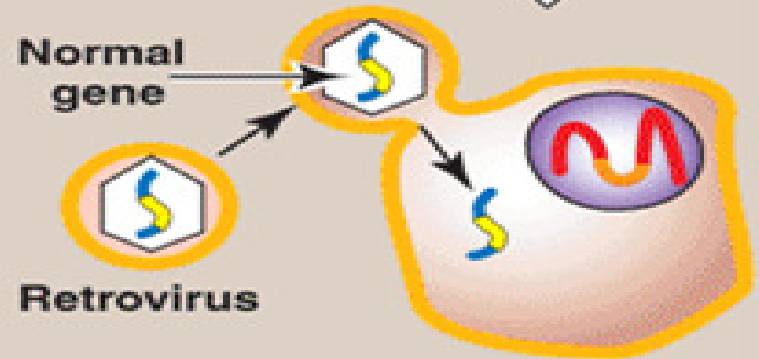
**1** Remove some of the patient's blood. Lymphocytes are separated from other cells in the blood)

**4** Return the cells to the patient, and the patient regains immune function



Defective gene

**2** Infect the lymphocytes with a retrovirus modified to carry the normal adenosine deaminase gene.



**3** The gene becomes integrated into the cell's chromosomes and is expressed. Some of the patient's lymphocytes now synthesize *adenosine deaminase*.

